

**FACILITATING RESEARCH WITH LEARNER DATA IN  
ONLINE SPEAKING TEST**

---

A Thesis

Presented to the

Faculty of

San Diego State University

---

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Computer Science

---

by

Rachana Santosh Bedekar

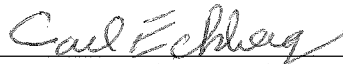
Fall 2011

**SAN DIEGO STATE UNIVERSITY**

The Undersigned Faculty Committee Approves the

Thesis of Rachana Santosh Bedekar:

Facilitating Research with Learner Data in Online Speaking Test



---

Carl Eckberg, Chair  
Department of Computer Science



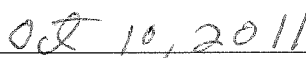
---

Joseph Lewis  
Department of Computer Science



---

Mary Ann Lyman-Hager  
Language Acquisition Resource Center



---

Approval Date

Copyright © 2011  
by  
Rachana Santosh Bedekar  
All Rights Reserved

## **DEDICATION**

I would like to dedicate this thesis to my father Santosh Bedekar and my mother Manali Bedekar for inculcating in me the importance of hard work and for supporting me always; to my sister Tanaya Bedekar for her encouragement and love.

## **ABSTRACT OF THE THESIS**

Facilitating Research with Learner Data in Online Speaking Test

by

Rachana Santosh Bedekar

Master of Science in Computer Science

San Diego State University, 2011

Language Acquisition Resource Center (LARC) is a Language Resource Center at San Diego State University. LARC has an online speaking test called CAST (Computer Assisted Screening Tool). CAST tests the speaking proficiency of the test takers in English, Spanish, Japanese, Hindi, French, Modern Standard Arabic, Persian, Mandarin Chinese, Iraqi Arabic, and Egyptian Arabic. For research purposes, a language corpus is created from the data obtained from the CAST test, the pre-survey and post survey taken before and after the test. A language researcher at LARC uses this language corpus to study various different factors that help language speakers achieve the Advance level proficiency. He also uses the language corpus to evaluate the reliability and validity of the CAST test, to facilitate the feedback loop that gives the test-takers meaningful assistance, to figure out the elements that constitute successful communication, to study the rater's rating strategies, and to study a test taker's attitude towards the computer based test.

The main objective of this thesis is to make language researchers' lives easy at LARC by providing them with an interface that allows them to efficiently search and navigate through the language learners' data. This data can be viewed as a web page and in the form of an Excel sheet. The web interface is an easy-to-use and consolidated outlook to the test data, pre-survey data, and post-survey data stored in the back-end. The research data can be sorted and filtered to develop data patterns and it can be modified and downloaded into a workable format as required. The researcher has access to the test ratings and recordings as a part of a quick menu. The research data is stored using Oracle databases; Java Servlets technology is used to congregate together the data and present it in a grid format; jQuery and JavaScript technologies are used to implement the sorting and filtering functionalities on the data table. A popup/context menu is created for the data table using JavaScript technology. The quick menu is used for sorting data, viewing the audio samples of the language learner and language learner information; jQuery is used to show/hide data columns in the grid as required generating data patterns and making it easy to browse.

## TABLE OF CONTENTS

	PAGE
ABSTRACT.....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
LIST OF ACRONYMS .....	x
ACKNOWLEDGMENTS .....	xi
<b>CHAPTER</b>	
1 INTRODUCTION .....	1
1.1 About LARC .....	1
1.2 About CAST .....	2
1.3 Need for Facilitating Research with Learner Data in CAST .....	2
2 COMPARISON WITH OTHER LANGUAGE LEARNING TOOLS .....	5
2.1 Other Language Research Tools Available .....	5
2.2 Learning from Research.....	7
3 REQUIREMENTS AND FUNCTIONAL SPECIFICATIONS.....	8
3.1 Functional Requirements .....	8
3.2 System Requirements.....	8
3.3 Technologies .....	9
3.3.1 Java Servlets.....	9
3.3.2 jQuery .....	9
3.3.3 JavaScript.....	10
3.3.4 JXL – JExcel API.....	10
4 SOFTWARE AND SYSTEM DESIGN AND DEVELOPMENT.....	11
4.1 Class Diagram .....	12
4.2 Activity Diagrams .....	12
4.3 Database Design Diagram.....	14
4.4 Application Screens .....	16
4.4.1 Excel Sheet Functionality and Show/Hide Functionality .....	16

4.4.2 Pre-Survey Excel Sheet.....	18
4.4.3 Post-Survey Excel Sheet.....	18
4.4.4 Column Show/Hide Functionality .....	18
4.4.5 View User's Information .....	21
4.4.6 The Popup Menu.....	21
4.4.7 Multiple Column Filters.....	24
5 TESTING.....	28
6 CONCLUSION AND FUTURE WORK .....	31
6.1 Conclusion .....	31
6.2 Future Work .....	31
BIBLIOGRAPHY .....	32
APPENDICES	
A DIAGNOSTIC ASSESSMENT USING CAST CONSENT FORM .....	33
B DIAGNOSTIC ASSESSMENT (EXTENSION OF CAST: RENEWAL APPROVAL FOR PROTOCOL #02-09-278).....	35
C USER MANUAL.....	39

**LIST OF TABLES**

	PAGE
Table 2.1. Comparing Language Proficiency Tools .....	6
Table 5.1. Testing Results.....	29



## LIST OF FIGURES

	PAGE
Figure 3.1. J2EE architecture.....	9
Figure 4.1. UML class diagram. ....	12
Figure 4.2. Activity diagram for the application.....	13
Figure 4.3. Part A of activity diagram in Figure 4.2--exporting data in Excel sheets. ....	14
Figure 4.4. Part B of activity diagram in Figure 4.2--pop up menu. ....	15
Figure 4.5. Part C of activity diagram in Figure 4.2--show/hide columns. ....	16
Figure 4.6. Database diagram for facilitating research with learner's data. ....	17
Figure 4.7. Main screen of the application for facilitating research at LARC.....	17
Figure 4.8. Data columns show/hide and Excel sheet report functionality. ....	18
Figure 4.9. Excel report of the pre-assessment survey data with global ratings.....	19
Figure 4.10. Post survey data with global ratings exported in Excel.....	19
Figure 4.11. Uncheck a category to hide it in the learner's data table.....	20
Figure 4.12. Uncheck TestID and name categories to hide the corresponding columns in the learner's data table . ....	20
Figure 4.13. Checked TestID and TestID column reappears.....	21
Figure 4.14. User's personal information visible to administrators.....	22
Figure 4.15. User's personal information access restrictions. ....	22
Figure 4.16. Pop up menu for users. ....	23
Figure 4.17. Test takers' information is anonymous to the users. ....	23
Figure 4.18. Quick pop up menu. ....	23
Figure 4.19. Menu to view learner's audio recordings and global ratings by selected rater for administrators.....	24
Figure 4.20. Selecting rater to view his ratings. ....	24
Figure 4.21. Review CAST ratings screen.....	25
Figure 4.22. Filtering data columns to narrow number of rows. ....	25
Figure 4.23. Filters column by one letter of the filter criteria.....	26
Figure 4.24. Filters column by two letters of filter criteria.....	26
Figure 4.25. Filters data columns by multiple filter criteria. ....	27

## **LIST OF ACRONYMS**

ACTFL – American Council of the Teaching of Foreign Languages

CAST – Computer Assisted Screening Tool

GUI – Graphical User Interface

IRB – Institutional Review Board

LARC – Language Acquisition Resource Center

LRC – Language Resource Center

LTI – Language Teaching Institute

## **ACKNOWLEDGEMENTS**

This thesis would have been a difficult undertaking without the direction of my thesis committee chair, Professor Carl Eckberg. I am also grateful to my other committee members, Professor Mary Ann Lyman-Hager and Professor Joseph Lewis for their constant support and guidance towards the development of this thesis. I would also like to thank Mike Pasamonik, Nada Novakovic and Trevor Shanklin for their help and support.

## **CHAPTER 1**

### **INTRODUCTION**

A language learning resource center promotes learning and teaching of foreign languages. It offers professional development workshops, generates language-learning materials, and conducts research on foreign language learning. Language Acquisition Resource Center (LARC) at San Diego State University (SDSU) is also a similar LRC that conducts research on foreign language learning.

The language researchers at LARC have to study a language corpus, which is a large collection of learner data to understand with precision how a learner performs at different difficulty levels and many other factors that help improve the quality of the CAST (Computer Assisted Screening Tool) test at LARC. Consequently, there was a need to create an application that will facilitate this research at LARC.

#### **1.1 ABOUT LARC**

In 1990, the first Language Resource Centers (LRCs) were established at U.S. universities by the Department of Education to address the increasing national need for expertise and proficiency in foreign languages. Now there are 15 LRCs to promote the teaching and learning of foreign languages. LRCs provide language learning and teaching materials to promote professional development opportunities for teachers and instructors, and carry out research on foreign language learning. Some LRCs concentrate on specific language areas and some on foreign languages in general, all have the common goal to develop to improve foreign language education in the United States [1].

LARC at SDSU is one of LRCs. LARC's mission is to develop and support the teaching and learning of foreign languages in the United States through research, technology, and publications. Particular attention is paid to less commonly taught languages, cross-cultural issues, language skills assessment, and teacher training [2].

## **1.2 ABOUT CAST**

CAST is one of the many programs hosted by LARC. CAST is used to test the oral proficiency of a test taker in a number of foreign languages. Examples of other projects would be the SPT at UT Austin. They have created a Spanish corpus and proficiency level the training. It will help in the assessment of a learner's level of proficiency and create a model for training of teachers in the application of language proficiency [3]. The intention of the CAST is to: (1) Elicit an adequate speech sample from examinees to assure a reliable floor (or baseline) rating. (2) Provide feedback to the examinees on how their proficiency can be improved. (3) Estimate examinees likely score on the official ACTFL Oral Proficiency Interview (OPI). (4) Provide positive feedback for proficiency-based teaching. With the newly added "reviewer" interface, instructors can effectively use the CAST as a tool to rate their students' oral production and to increase their own understanding of the criteria by which ACTFL ratings are given [4].

With regard to CAST, it is the intersection of three fields: computational linguistics, applied linguistics (corpora) and language pedagogy. A researcher would have his/her hands full in any one of these fields. The researcher may not end up helping the test taker, but the field as a whole, or the instructor. Perhaps where CAST best shows the intersection of research and benefits for the test-taker are the way the feedback cycle is set up. A language corpus is created of spoken learner data from the CAST, which is helpful for the following:

- An evaluation of the reliability and validity of the test itself.
- Facilitating the feedback loop so that the test-taker gets meaningful assistance.
- In the analysis of the elements that constitute successful communication.
- To study rater's strategies.

## **1.3 NEED FOR FACILITATING RESEARCH WITH LEARNER DATA IN CAST**

A person may be not yet ready to take the full blown OPI. There is no direct feedback from ACTFL. This is an attempt to provide test-takers feedback. There are several levels defined to indicate the proficiency of the speaker: "Advance level", "At level", "Not at Level". ACFTL works together with CAST and NCATE for training teacher to understand about the proficiency levels.

For a foreign language researcher, just looking at a single score of pass or no pass at an audio test is not useful enough in terms of understanding to contribute to achieving advance level, so it is essential to find out what factors or individual profiles like the background, education, etc. combine together that contribute in achieving Advance level. Ray T. Clifford, LARC board member and Associate Dean (Director, Center for Language Studies) rightly said, “Life begins at the Advance Level.”

It is a part of the language researcher’s job to determine how a person who has attained an advance level handles a situation. In addition, what are the factors that facilitate achieving an Advance level?

I interviewed Dr. Mary Ann Lyman Hager (Director of LARC), Nadezda Novakovic (Researcher at LARC), Dr. Trevor Shanklin (Director of Lab Operations, LARC) and Mike Pasamonik (Software Engineer) as a part of requirements gathering. Dr. Mary Ann Lyman Hager had in mind a project idea for creating an application for language proficiency data that can be used by the researchers at LARC for various different research purposes. She introduced me to Dr. Trevor Shanklin and Nadezda to collect their requirements and Michael Pasamonik to know more about the CAST software system.

When interviewed, Dr. Trevor Shanklin, Director of Lab Operations, LARC, said that a researcher at LARC is responsible for studying a language corpus created using the CAST test data and involves at least one of the following:

1. An evaluation of the reliability and validity of the test itself.
2. Facilitating the feedback loop so that the test-taker gets meaningful assistance.
3. It will help in the analysis of the elements that constitute successful communication.

He wanted to see a test-takers demographic data like name, age, gender, country of residence, etc. along with other test data. This is useful in evaluating an individual by comparing him with heritage speakers.

When interviewed, Nadezda Novakovic thought it would be great to have research data presented in a coordinated formatted. She also thought it would be a good idea to link the post-assessment survey data to the candidates’ final CAST ratings. She came up with the following additional requirements for this project:

1. A separate grid of responses to the pre-assessment survey with a column containing candidates’ CAST ratings.

2. A separate grid of responses to the post-assessment survey, again with a column containing candidates' CAST ratings.
3. A big grid containing the responses to both the pre-assessment and post-assessment surveys with a column containing CAST ratings.

I interviewed Mike Pasamonik, to understand the CAST system from a software engineer's perspective. This project is incorporated as a CAST module, so it had to be compatible with the current CAST structure. He explained to me the different technologies used to create the CAST system, database structure, server requirements, hardware requirements, and software requirements.

Keeping all these things in mind, it was required to develop a system that will have interfaces and a database to organize, store and display the different types of research data in a comprehensible web-based view. To make sure the confidentiality issue of the test-taker is not ignored, the personal data is visible only to the Administrators and not to other users/researchers. LARC has obtained permission from the IRB to use the test-taker's data; the permission document is included in Appendix A and B.

## **CHAPTER 2**

### **COMPARISON WITH OTHER LANGUAGE LEARNING TOOLS**

This section discusses a few language proficiency research programs, how they carry out language research and how the research was useful to make the LARC language proficiency research application helpful for researchers.

#### **2.1 OTHER LANGUAGE RESEARCH TOOLS AVAILABLE**

CALSYS at the University of Oregon is also a Language Resource Center with a goal to improve the teaching and learning of foreign languages [5]. CALSYS has a searchable database software called SLAWeb. SLAWeb is a searchable database of spoken and written productions. The data of the database is composed of learners in grades 6-16 by using an online proficiency assessment test called Standards-Based Measurement of Proficiency (STAMP) [6].

The SLAWeb tool has research productions from more learners of Chinese, French, German, Hebrew, Japanese, Spanish, and Turkish. The CASLS databank can be searched using different parameters like school level, class level, immersion and heritage status, and proficiency [6].

The research on SLAWeb was useful to see what others are doing in the field of language proficiency research, what are the pros and cons of the software and how can this analysis be used to improve the LARC language proficiency research. In case of LARC, data is collected from language learners who take the CAST test which is an online proficiency assessment. The data is from learners of Persian, ESL, Spanish, Testing, Chinese, Egyptian, Iraqi, French, Pashto, German, Hindi, Tagalog, Urdu, Bengali, Japanese. Therefore, there was a need for this thesis project to assist a researcher in understanding with precision, how a learner performs at different test levels, and come up with ways to improve feedback to the test takers. Table 2.1 [4], [5] represents the comparison between LARC's learners' database application and CALSYS's learners' database application.



**Table 2.1. Comparing Language Proficiency Tools**

	<b>LARC</b>	<b>CALSYS (SLAWeb)</b>
<b>Data collected from</b>	CAST is LARC's online speaking test and the post and pre survey carried out during the test.	An online proficiency assessment, the Standards-Based Measurement of Proficiency (STAMP).
<b>Type of test data</b>	Includes speaking results and assessment survey data.	Includes reading, writing, and speaking results.
<b>Shows test results</b>	By demographic data such as age, gender, country of residence and years of study, primary education.	By language of study, native language, and grade and demographic data such as age and years of study.
<b>Information Security</b>	Login by user name and password.	Login using username and password.
<b>Searchable by</b>	A number of parameters, including school level, class level, proficiency level, and heritage status.	Parameters, including school level, class level, proficiency level, and immersion or heritage status.

Sources: Language Acquisition Resource Center. Computer Assisted Screening Tool, n.d. <http://cast.sdsu.edu/>, accessed Sept. 27, 2011; University of Oregon. Searchable Proficiency Database, n.d. <http://casls.uoregon.edu/pages/research/proficiencydatabase.php>, accessed Sept. 27, 2011.

The Harvard University Department of Linguistics uses Data coding primarily done with Excel/OpenOffice or makes use of software called PXLab. They can conduct simple experiments, but no (official) support is available when they encounter any problems [7].

The intent of the research was to find out an ideal way to preserve the privacy of the learners and cause no harm to them. A method was to be selected to implement these criteria keeping in mind the consistency with the existing system.

While creating the corpus from the test-takers' data, the confidentiality issue of the test-taker's data is given utmost importance and it is done under a lot of protection. According to Talk Bank, The British Association for Applied Linguistics, LSA Ethics Discussion, and Confidentiality and Informed consent, it is a responsibility of linguistics individually and collectively to anticipate ethical dilemmas and avoid harm to those with whom they work. According to the code of ethics, no personal data about individuals is to be made available to the users apart from audio records and basic facts [8], [9], [10], and [11].

According to Talk Bank and British Association, the group of individuals involved in this research is classified into three groups [8], [9]:

1. Participants – people who have been recorded
2. Contributors – researchers who have collected recordings
3. Users – researchers who receive data from TalkBank.

The LARC has a similar classification for the users for learners' data application and additionally a fourth category called administrators. This classification is used when giving administrators or users access to learner's data. The administrators can see a learner's contact details, in case they want to contact the learner for reasons like retaking the exam and reporting progress or retreat.

There are many different ways practiced by different organizations to protect the confidentiality of the participants or informants. The options that can be implemented at LARC were as follows:

1. Determine in advance whether the test taker or participant wants to remain anonymous or receive recognition.
2. Restrict access to such data to bona fide users.
3. Make data at the data entry level anonymous; to make sure it does not cause loss of data.

In a meeting with Dr. Mary Ann Lyman-Hager, it was decided that the second option would be selected, as the other two options would require changes to the current flow of the CAST system.

## **2.2 LEARNING FROM RESEARCH**

The technique adopted in this application, is to restrict access to such data to authorized users and make it anonymous to all the users except administrators. A robust system has been set up in the application to make the personal data anonymous when required.

LARC and the users need to follow a code of ethics while working on the test-takers' data. LARC has to assure that users respect the rights of participants. The IRB at the university reviews all research conducted at the institution. SDSU also has an IRB. LARC has a document from IRB (see Appendix A and B).

## **CHAPTER 3**

### **REQUIREMENTS AND FUNCTIONAL SPECIFICATIONS**

This section discusses the necessary software resources and hardware components required for the Learner's database system to run effectively. Further, it covers the behaviors and properties required by its users and other components of this application.

#### **3.1 FUNCTIONAL REQUIREMENTS**

This section describes the functional requirements of this application:

- Create a user-friendly interface for the learner's data.
- Preserve confidentiality of test-taker or learner by limiting access to data.
- Create Excel sheets for the pre-survey, post-survey and both.
- The learner's data should be easy to sort and filter.
- The filter can be applied to one to many categories at the same time.
- Limit the data to be displayed on the screen by giving user and option to select the required data.

#### **3.2 SYSTEM REQUIREMENTS**

The details of the development environment for the CAST are given below. The hardware environment:

- Xserve PowerPC G5 (Type: 64Bit OS and Speed: 2.0Ghz)
- 2 GB RAM
- 153 GB Hard Disk

In addition, the software environment:

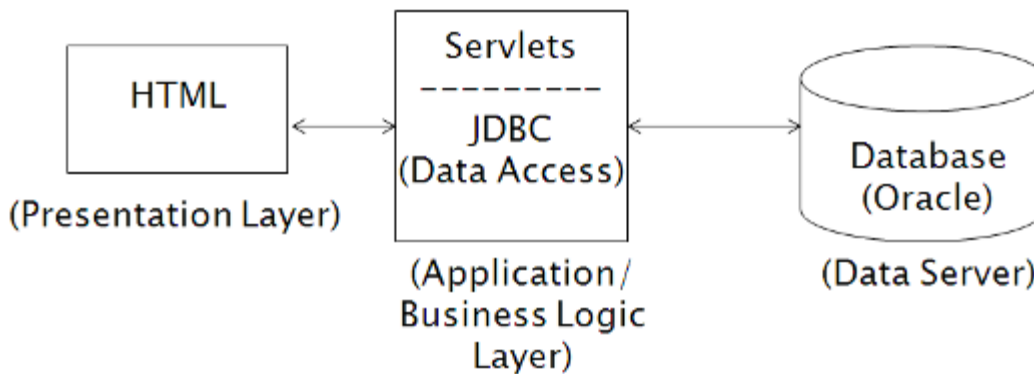
- Mac OS X Serve v 10.4.11
- Java Standard Edition v 1.5 SDK
- Orion
- Oracle 10g Database
- Any browser with JavaScript enabled

### 3.3 TECHNOLOGIES

This section describes the programming technologies used to build the learner's data application.

#### 3.3.1 Java Servlets

CAST test is written using Java 2 Platform (J2EE) programming standard. A servlet is a Java programming language class used to expand the capacity of servers that host applications accessed via a request-response programming model [12]. This application is included as a module of the CAST application as a whole and as a result, the same programming language standards were used for this application as well. This application is built using the J2EE architecture as shown in the Figure 3.1.



**Figure 3.1. J2EE architecture.**

#### 3.3.2 jQuery

jQuery is a fast and concise JavaScript Library that simplifies HTML document traversing, event handling, animating, and Ajax interactions for rapid web development. jQuery is designed to change the way that you write JavaScript [13]. jQuery code can be easily embedded in HTML; as a result, jQuery was a choice for implementing show/hide data functionality as HTML is used for front-end coding for this application.

jQuery was used to map the checkboxes with the data table. This functionality can also be implemented using JavaScript, but jQuery reduces the amount of code and also is easier to maintain compared to JavaScript. Moreover, some people disabled JavaScript but this is not done for jQuery.

### **3.3.3 JavaScript**

JavaScript is a lightweight programming language. JavaScript is generally embedded directly into HTML pages or imported as a “.js” file. JavaScript is an interpreted language and as a result it is fast. JavaScript is useful to respond to user events like click, mouse in, mouse out, etc on the interface.

This application has a popup menu or context menu to link to audio recordings and view learners’ information. The context menu is a div element with display property set to none and position property set to absolute. The context menu appears on the mouse down event on the learners’ data table. This context menu works with all majors browsers like IE, Firefox, Chrome, and Opera [14].

### **3.3.4 JXL – JExcel API**

Java Excel (JExcel) API is an open source Java API that facilitates developers to read, write, and modify Excel spreadsheets dynamically and write the changes to any output stream [15].

There was a requirement that the code of this application should be compatible with the code structure of the CAST test, which is written in Java using Java Servlets. The JExcel API can be invoked from within a servlet, thus giving access to Excel spreadsheets over the internet and intranet web applications. As a result, JExcel API was used as it can be used with Java Servlets to implement the export data component of this application.

## CHAPTER 4

### SOFTWARE AND SYSTEM DESIGN AND DEVELOPMENT

In designing such an application, there exist two possible approaches: procedural and object-oriented. The procedural approach involves creating algorithmic procedures then breaking the necessary steps into modules through which data is input logically. Whereas, the object-oriented approach focuses on terms called objects. Objects encapsulate both procedures and data in a single entity. There also exists an 'object interface', which defines how the object interacts with the rest of the code.

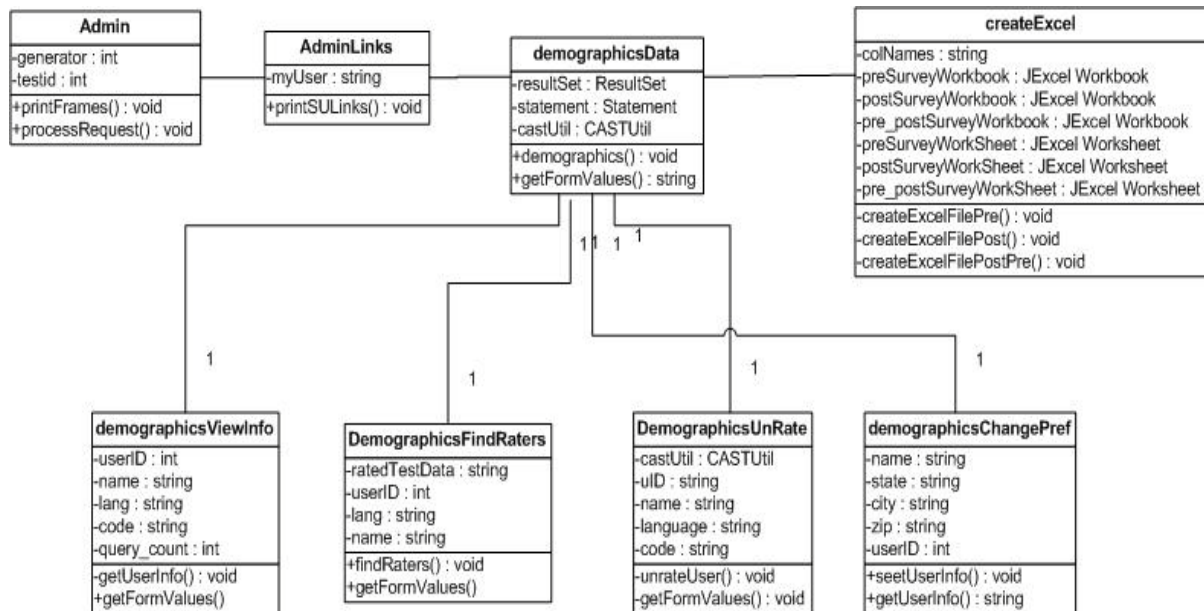
It was a part of the requirements that this application should be compatible with the current CAST application. The CAST application is created using Java Servlets thus an object-oriented design methodology was selected. In addition, structured query language (SQL) provided a database query language. JDBC communicates SQL to Oracle.

The Unified Modeling Language (UML) based approach was used design the interface and database components of this project. UML design uses four types of structures to facilitate language-independent organization, visualization, and creation of an object-oriented program:

- *Class Diagram:* It depicts system organization by showing the system's classes, relationships between the classes, and class attributes and behaviors (procedures).
- *Database Diagram:* It is an organization of data to represent an outline of how the database will be constructed and depicts the relationship between the different data tables.
- *Activity Diagrams:* A state-transition diagram, which depicts the consecutive transitions of an object from various states during execution of a given procedure. A state represents the conditions of the object as the program executes. The transitions begin at an initial state and progress either in parallel or conditionally to a final state.
- *Sequence Diagrams:* A chronological diagram, which depicts the sequence of actions that occur in a system, including the invocation of object methods, modification of object properties, and input/output between the user, interface, and database.

## 4.1 CLASS DIAGRAM

The class diagram for this application (Figure 4.1) shows the structure and interaction of the classes. Each entity class (shown boxed in Figure 4.1) models the properties and behaviors of the data. For example, the class for Demographics Data will contain properties such as resultSet, and statement, and the behaviors of +demographics (), +getFormValues (), etc. The lines represent the relationship between the classes.

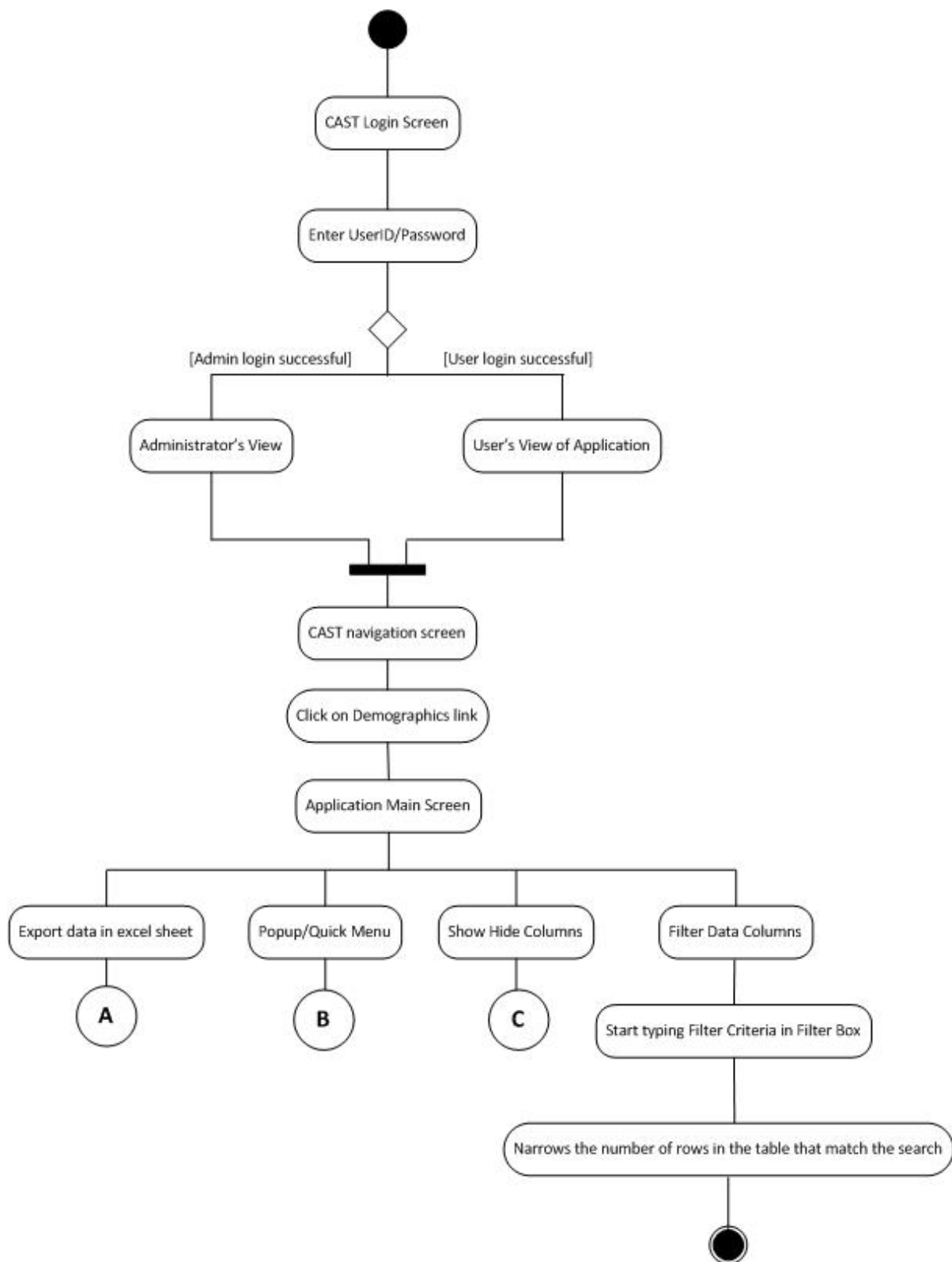


**Figure 4.1. UML class diagram.**

## 4.2 ACTIVITY DIAGRAMS

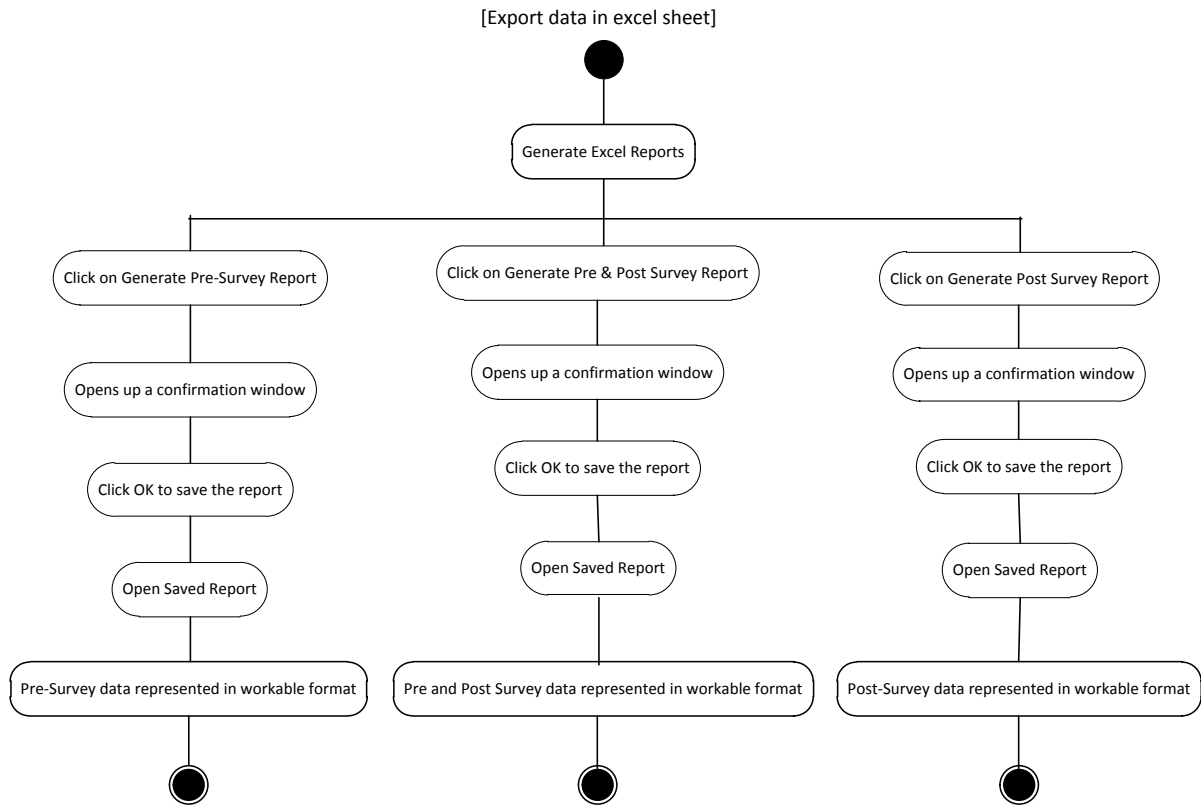
The activity diagram in Figure 4.2 demonstrates the logical flow of actions in this application when accessing the data contained in objects of its classes. If the user wants to access the proficiency database, he has to login using his user ID and password. If he is an administrator he will have a different view of the application as compared to the other users. He can then access the application by clicking on the Demographics link. The user will then have four different choices like (1) export data in Excel sheet, (2) show/ hide data columns to see only required columns, (3) use the Popup menu to access test-takers information, (4) filter the columns to narrow the number of rows visible. As not all the activities can be depicted in one diagram, the diagram is split into three different sections A, B, and C.

Figure 4.3 depicts the Excel functionality of the application and section A of Figure 4.2 in detail. The user can generate Pre-Assessment Survey report with global ratings,



**Figure 4.2. Activity diagram for the application.**





**Figure 4.3. Part A of activity diagram in Figure 4.2--exporting data in Excel sheets.**

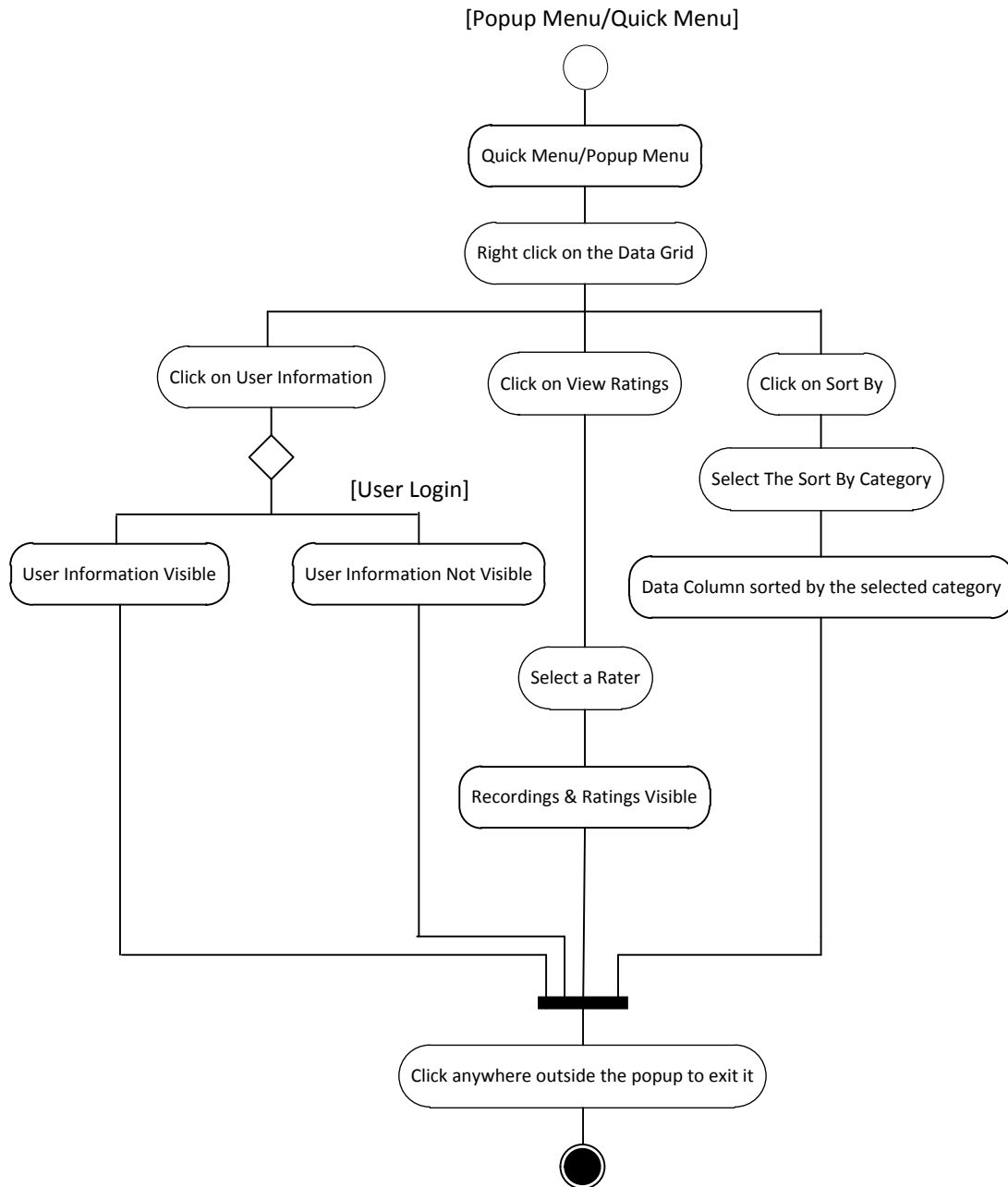
Post-Assessment Survey report with global ratings, Pre, and Post-Assessment Survey report with global ratings.

Figure 4.4 depicts the activity diagram for the pop-up menu functionality and section B of Figure 4.2. The pop-up menu gives the user three different options like: (1) Sort the records by Name, ID, and Language, (2) View User's Information, and (3) View the Audio recordings and ratings of the user.

Figure 4.5 is the activity diagram is an extension of section C of Figure 4.2. The application has the capability to limit the data the user wants to see on the screen, as there is a large amount of research at one place.

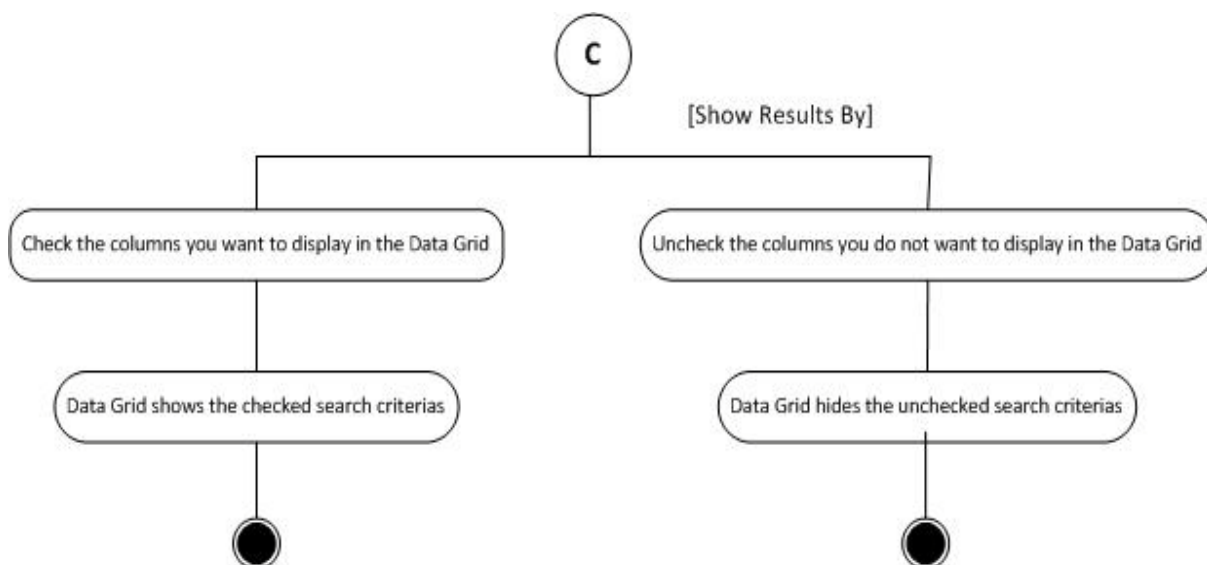
### 4.3 DATABASE DESIGN DIAGRAM

SQL serves as an interface to a Relational Database Management System (RDBMS), which manages data organized in the form of related tables. This feature preserves security and authentication, and provides powerful database-driven functionality.



**Figure 4.4. Part B of activity diagram in Figure 4.2--pop up menu.**

Each table represents a class entity, i.e., CASTPRESURVEY\_ANSWERS, CASTPOST\_ANSWERS, RATEDTESTS, DEMOGRAPHICS\_GROUP, and BETAREG along with its attributes and its links, representing its relationships to other entities. Two entities relate to each other in the following ways: Primary Key (PK) and Foreign Key (FK). For each entity, the PK defines a field unique to each instance of the Activity Diagram Part B--pop up menu and activity Diagram part C--show/hide data columns entity, helping to



**Figure 4.5. Part C of activity diagram in Figure 4.2--show/hide columns.**

identify the instance with respect to other instances. Figure 4.6 represents the PK in bold at the top of each entity table. Another entity may manipulate a given instance by accessing the instance's PK. In such case the PK referred to as the FK of the manipulating entity, identifies the instance. For example, USER ID in BETAREG represents a suitable PK, since no two identical User IDs exist in the database. This key appears as the FK in CASTPRESURVEY\_ANSWERS, CASTPOST\_ANSWERS, RATEDTESTS, and DEMOGRAPHICS\_GROUP. Any database query fired on an entity would reflect the changes in that entity as well as any other dependant entity. Therefore any change to the PK of BETAREG, will also change the attribute values of the related entities.

## 4.4 APPLICATION SCREENS

This section gives the details of the GUI used to give a web view to the learner's data along with its important components.

### 4.4.1 Excel Sheet Functionality and Show/Hide Functionality

Figure 4.7 depicts the main screen of the application, which shows all the functionalities of the application in brief. Figure 4.8 screen shot zooms in to show the Excel sheet functionality and the show/hide column functionality. There are three links as follows:

1. CAST Pre-Survey Data - This Excel sheet contains all the CAST's Pre-Survey data.

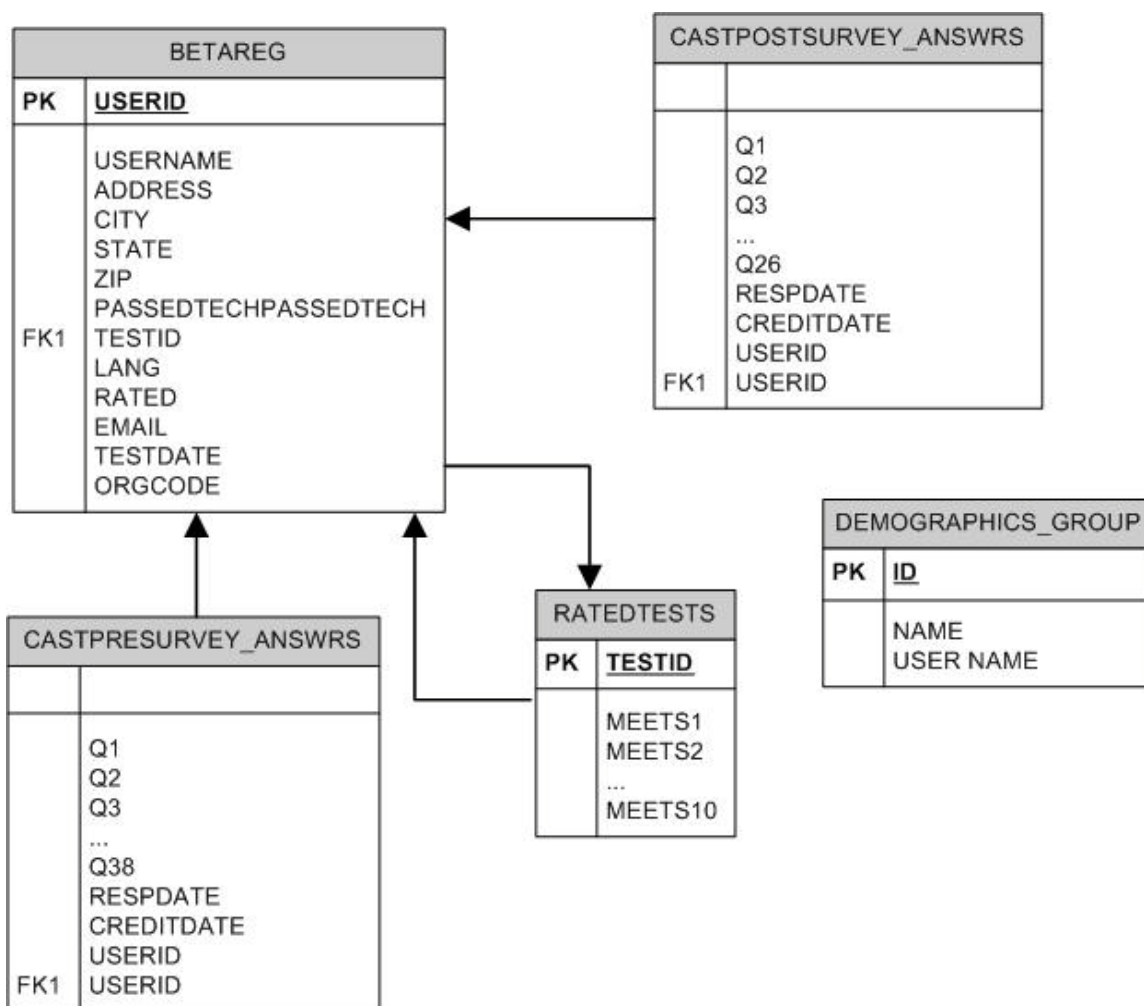


Figure 4.6. Database diagram for facilitating research with learner's data.

CAST Administration

Demographics

User Logged in: v\_jshah

Create Excel Sheet For: CAST Pre-Survey Data | CAST Post-Survey Data | CAST Post-Survey Data & CAST Pre-Survey Data

Select Columns To Be Displayed In The Table:

CAST Pre-Survey

- ☒ Test ID
- ☒ Name
- ☒ Language
- ☒ Test Date
- ☒ Registration Code
- ☒ Age
- ☒ Gender
- ☒ Language often spoken in your home?
- ☒ Your native language?
- ☒ Live in area - language was predominant
- ☒ Language of instruction (primary)
- ☒ Language of instruction (secondary)
- ☒ Taken a "Study Abroad" course
- ☒ Highest level of education
- ☒ Global Ratings

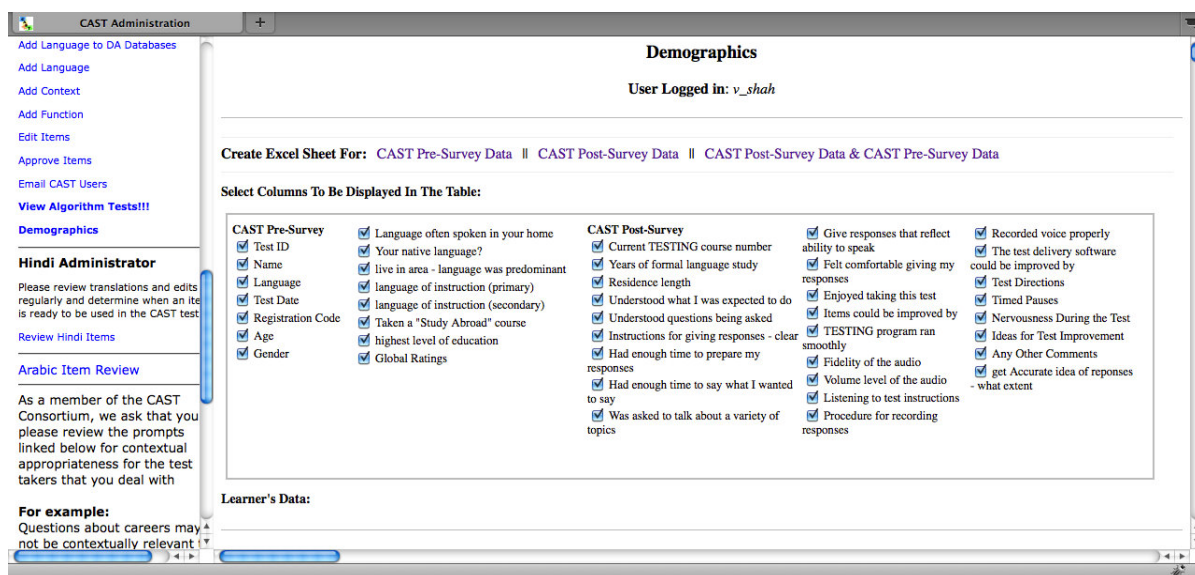
CAST Post-Survey

- ☒ Current TESTING course number
- ☒ Years of formal language study
- ☒ Residence length
- ☒ Understood what I was expected to do
- ☒ Understood questions being asked
- ☒ Instructions for giving responses - clear
- ☒ Had enough time to prepare my responses
- ☒ Had enough time to say what I wanted to say
- ☒ Was asked to talk about a variety of topics
- ☒ Give responses that reflect ability to speak
- ☒ Felt comfortable giving my responses
- ☒ Enjoyed taking this test
- ☒ Items could be improved by TESTING program ran smoothly
- ☒ Fidelity of the audio
- ☒ Volume level of the audio
- ☒ Listening to test instructions
- ☒ Procedure for recording responses
- ☒ Recorded voice properly
- ☒ The text delivery software could be improved by
- ☒ Test Directions
- ☒ Timed Pauses
- ☒ Nervousness During the Test
- ☒ Ideas for Test Improvement
- ☒ Any Other Comments
- ☒ get Accurate idea of responses - what extent

Learner's Data:

Test ID	Name	Language	Test Date	Registration Code	Age	Gender	Is language often spoken in your home?	Is it your native language?	Have you lived in an area where a different language was predominant for six months or longer?	In
pCCVAr	Jonathan Williams	es	2007-03-23 00:00:00	BETA_OPI	22	Male	No	English	No	Eng

Figure 4.7. Main screen of the application for facilitating research at LARC.



**Figure 4.8. Data columns show/hide and Excel sheet report functionality.**

2. CAST Post-Survey Data - This Excel sheet contains all the CAST's Post-Survey with global ratings. It is used by the researchers to compare the test-takers' mind-set by after the test and test results by viewing their global results and evaluation their post-survey answers at the same time.
3. CAST Post-Survey and CAST Pre-Survey - This Excel sheet contains all the CAST'S Post-Survey as well as Pre-Survey data with the global ratings. Each link is responsible for creating an Excel sheet. This gives the researcher an easy way to download the data in a workable format.

#### 4.4.2 Pre-Survey Excel Sheet

Figure 4.9 shows and Excel sheet of the Pre-Survey Data and the Global Rating of different test-takers. This Excel sheet gives researcher a workable format for the data.

#### 4.4.3 Post-Survey Excel Sheet

Figure 4.10 shows an Excel sheet of the Pre-Survey Data and the Global Rating of different test-takers. This Excel sheet gives researcher a workable format for the data. Similarly, an Excel sheet can be generated for Post-Survey and Pre-Survey both.

#### 4.4.4 Column Show/Hide Functionality

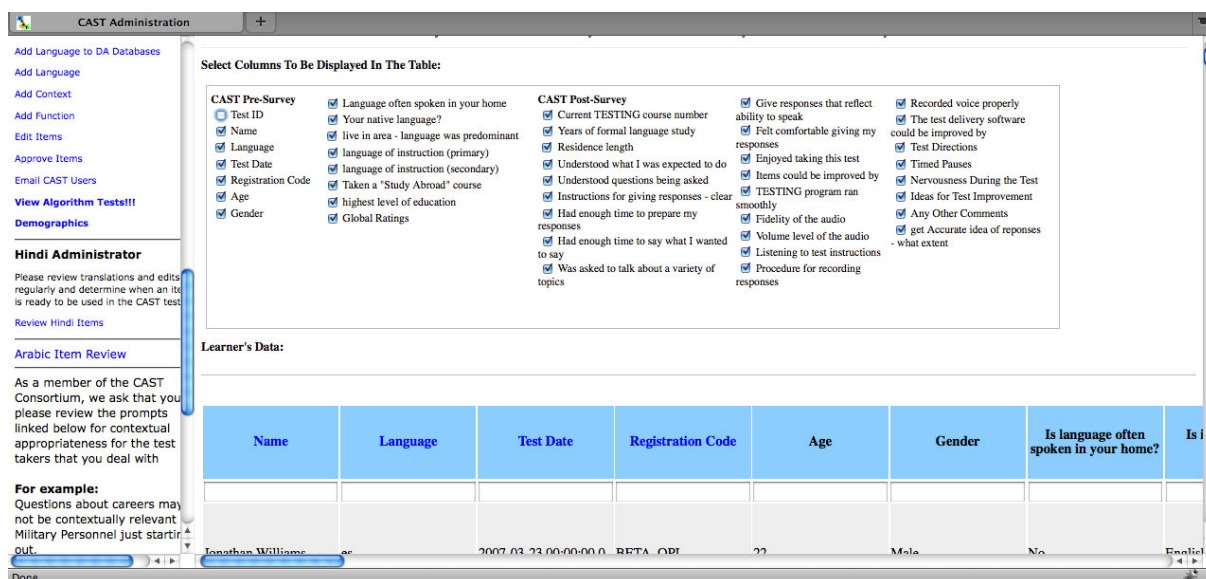
As there is a large amount of learner data, it is required to have a method to facilitate the researcher with functionality so that he/she can select only the required data columns they want instead of scrolling through the entire page horizontally. As shown in Figures 4.9 and

[illegible]

4.10, the checkboxes are connected to the corresponding columns and when one checks or un-checks a particular check box, its corresponding column is shown or hidden respectively.

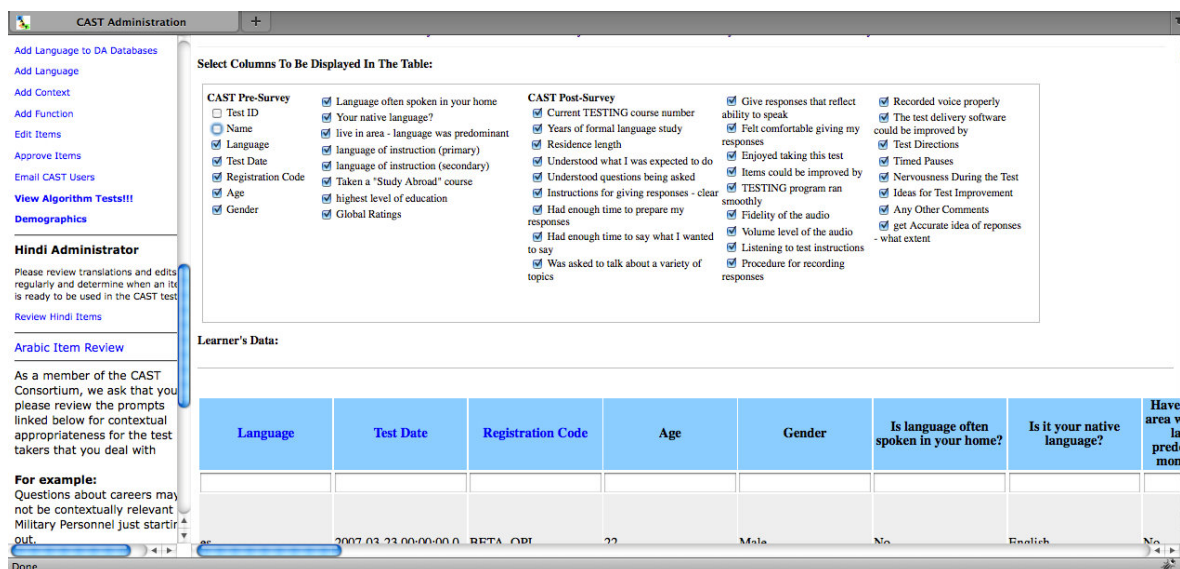
In Figure 4.11, the Test ID checkbox is unchecked and this hides the corresponding Test ID column in the Learner's Data Table.





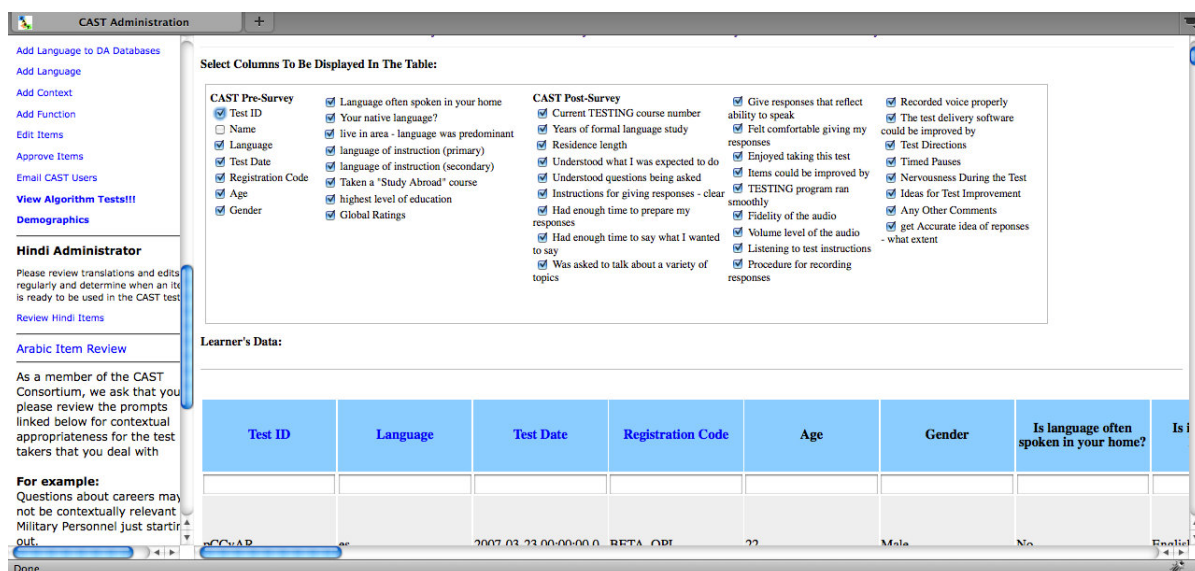
**Figure 4.11. Uncheck a category to hide it in the learner's data table.**

In Figure 4.12 the Name checkbox is unchecked and this hides the corresponding Name column in the Learner's Data Table.



**Figure 4.12. Uncheck TestID and name categories to hide the corresponding columns in the learner's data table .**

In Figure 4.13 the Test ID checkbox is checked again and this display the corresponding Test ID column in the Learner's Data Table keeping the Name column is still hidden. Therefore, the user can randomly select the columns he/she wants to see by checking their corresponding checkboxes.



**Figure 4.13. Checked TestID and TestID column reappears.**

### 4.4.5 View User's Information

For users other than the administrators, test takers' contact details are not visible, as it is considered a breach of confidentiality. When users other than administrators try to access the data, they will see it as anonymous. However, when the administrators right clicks on the learner's data table, as shown in Figures 4.14 and 4.15, they will be able to see the Test-Taker's contact details. Different similar applications have used a variety of practices as follows:

1. Make the data anonymous while entering it into the system.
2. Use login functionality to limit access to data.

However, the anonymity of data seems to be a very fine practice. At LARC, we have used the second method as listed above to protect data access. The language learners' data in the Figure 4.14 and Figure 4.15 are removed for confidentiality reasons.

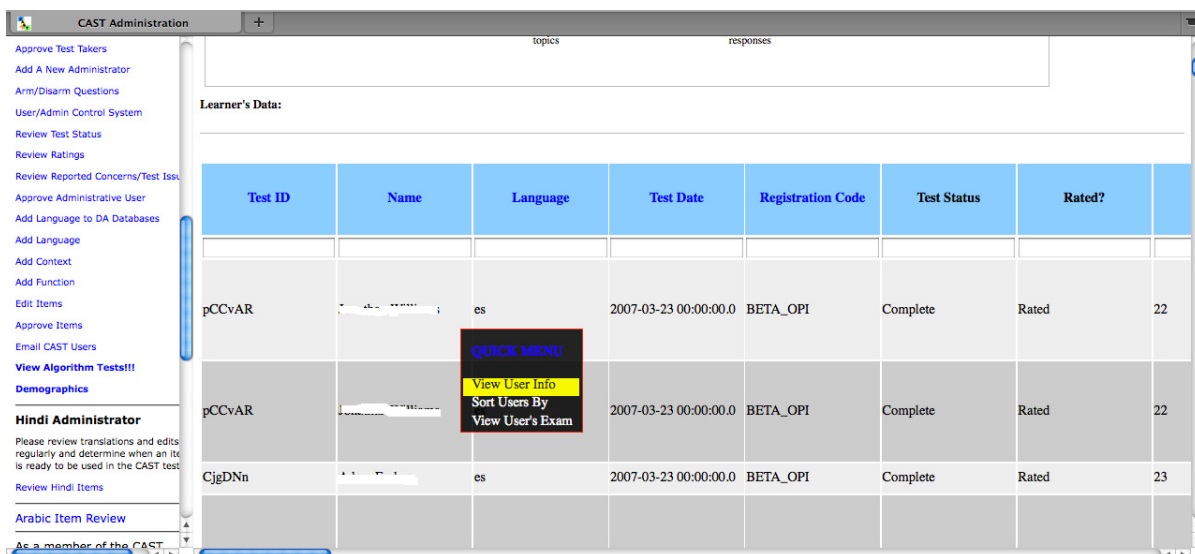
When users click on the View User Info as shown in Figure 4.16, they will not be able to see the test takers' personal information and all the information will appear as anonymous as shown in Figure 4.17.

### 4.4.6 The Popup Menu

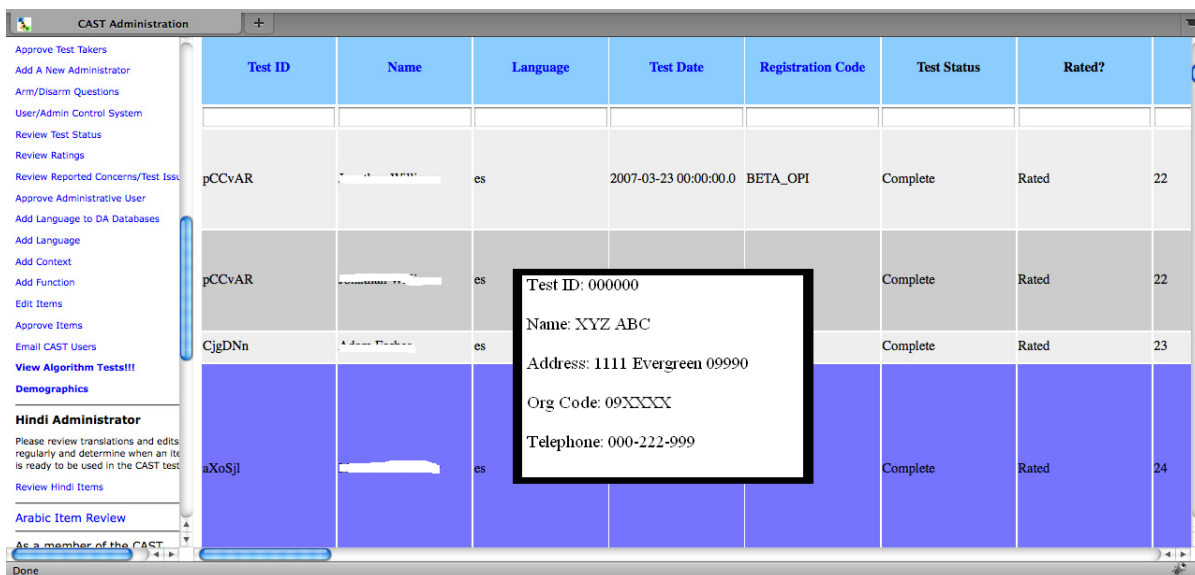
Figure 4.18 depicts the pop up menu created using JavaScript. It facilitates access to the test-takers' oral recordings.

When a user clicks on any record, a menu pops up with sub-menus like:





**Figure 4.14. User's personal information visible to administrators.**



**Figure 4.15. User's personal information access restrictions.**

1. View User's Exam – View Users' Exam Audio recordings and rater's ratings.
2. Sort Users By – When you click on this option the User can sort the data by the selected Category
3. View User's Info – This Information is visible only to Authorized users.

As shown in Figure 4.19, when the user uses the right click event on the Learner's data grid, a quick menu pops up. This menu has a "View User's Exam" option, it gives another "Rated By" list. Researcher can then select a rated exam by a particular rater as shown in Figure 4.20. Once the user clicks on the Rater's name it gives you another screen

The screenshot shows the 'CAST Administration' window. On the left is a sidebar with various administrative links. The main area displays a table of test takers under the heading 'Learner's Data:'. A 'QUICK MENU' pop-up is overlaid on the table, containing the following options:

- View User Info
- Sort Users By
- View User's Exam

Test ID	Name	Language	Test Date	Registration Code	Test Status	Rated?	
pCCvAR	Anonymous	es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22
pCCvAR	Anonymous		2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22
CjgDNn	Anonymous	es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	23

Figure 4.16. Pop up menu for users.

The screenshot shows the 'CAST Administration' window with the same table of test takers. A 'View User Info' pop-up is overlaid, displaying the following information:

- Test ID: Anonymous
- Name: Anonymous
- Address: Anonymous, Anonymous Anonymous Anonymous
- Language: Anonymous
- E-mail: Anonymous
- Test Date: Anonymous
- Org Code: Anonymous
- Test Status: Complete
- Rating Status: Rated

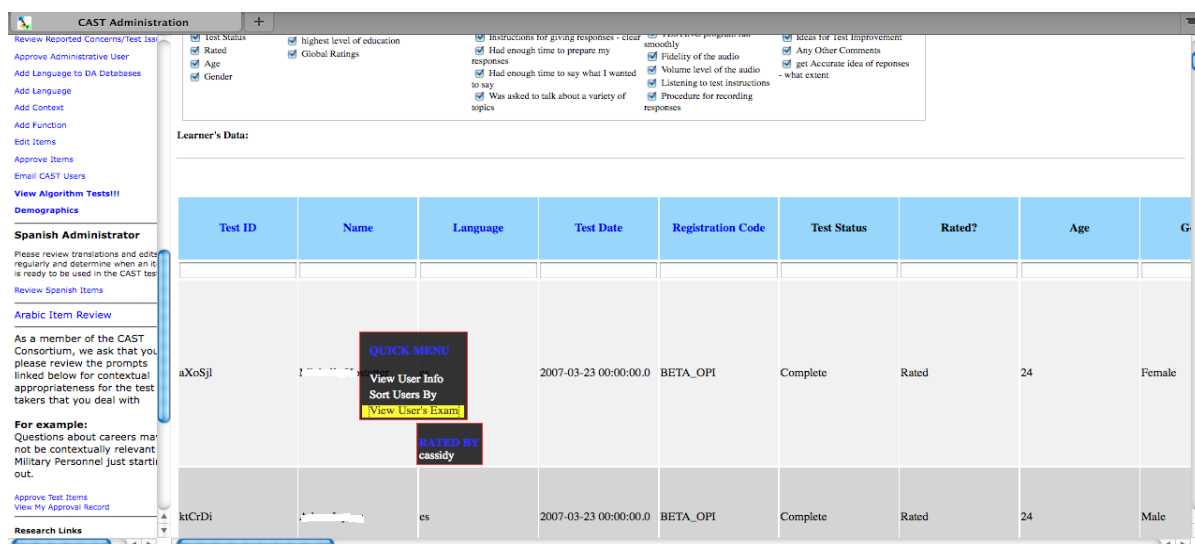
Test ID	Name	Language	Test Date	Registration Code	Test Status	Rated?	
pCvAR	Anonymous	es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22
pCvAR	Anonymous	es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22
gDNn	Anonymous				Complete	Rated	23
oSjl	Anonymous				Complete	Rated	24

Figure 4.17. Test takers' information is anonymous to the users.

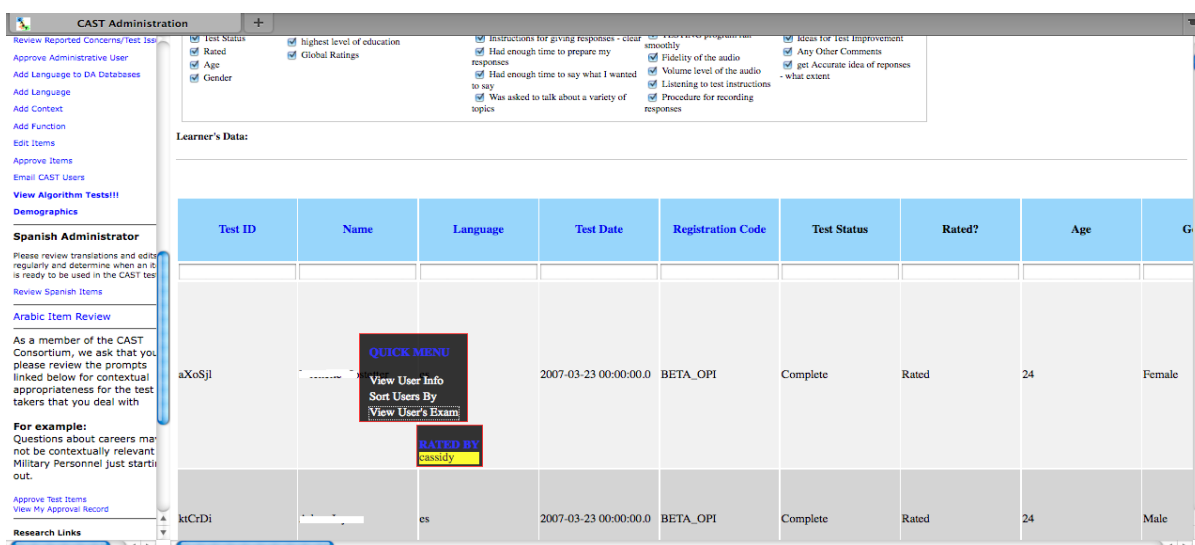
The 'QUICK MENU' pop-up contains the following options:

- View User Info
- Sort Users By
- View User's Exam

Figure 4.18.  
Quick pop up  
menu.



**Figure 4.19. Menu to view learner's audio recordings and global ratings by selected rater for administrators.**



**Figure 4.20. Selecting rater to view his ratings.**

that shows the test-taker's recordings and ratings by that particular rater as shown in Figure 4.21.

### 4.4.7 Multiple Column Filters

Figure 4.22 shows the method to start filtering data columns by typing filter criteria in the filter box above each data column. The multiple column filter feature enables the researcher to filter the data using more than one category. For instance, as shown in the Figures 4.23, 4.24 and 4.25.

Test ID: aXoSjl				
Question	Response	Prompt	Evaluation	Comments
1072_es	<a href="#">Download This Response</a>	<a href="#">View Test Item</a>	Meets Task Requirements	No Comments
1207_es	<a href="#">Download This Response</a>	<a href="#">View Test Item</a>	Meets Task Requirements	No Comments
1106_es	<a href="#">Download This Response</a>	<a href="#">View Test Item</a>	Meets Task Requirements	No Comments
1135_es	<a href="#">Download This Response</a>	<a href="#">View Test Item</a>	Meets Task Requirements	No Comments
1071_es	<a href="#">Download This Response</a>	<a href="#">View Test Item</a>	Meets Task Requirements	No Comments
<b>Overall Comments:</b>			No Comments	
<b>Global Rating:</b>			Full	

Figure 4.21. Review CAST ratings screen.

<div> <div> <div>CAST Administration</div> <div>+</div> </div> <div> <div> <div>Approve Items</div> <div>Email CAST Users</div> <div>View Algorithm Tests/II</div> <div>Demographics</div> </div> <div> <div>Hindi Administrator</div> <div>Please review translations and edit regularly and determine when an item is ready to be used in the CAST test.</div> <div>Review Hindi Items</div> <div>Arabic Item Review</div> </div> <div> <div>As a member of the CAST Consortium, we ask that you please review the prompts linked below for contextual appropriateness for the test takers that you deal with.</div> <div>For example: Questions about careers may not be contextually relevant to Military Personnel just starting out.</div> <div> <div>Approve Test Items</div> <div>View My Approval Record</div> </div> </div> <div> <div>Research Links</div> <div>View CAST Ratings</div> <div>View CAST Survey Results</div> <div>View Recording Scripts</div> <div>View CAST Function/Context Approval Grid</div> </div> <div> <div>Hindi Instructor Review</div> <div>4</div> <div>1</div> </div> </div> <div> <div> <div>to say</div> <div> <input checked="" type="checkbox"/> Was asked to talk about a variety of topics </div> </div> <div> <div>Listening to new instructions</div> <div> <input checked="" type="checkbox"/> Procedure for recording responses </div> </div> </div> </div>									
Learner's Data:									
Language	Test Date	Registration Code	Test Status	Rated?	Age	Gender	Is language often spoken in your home?	Is it your language?	
es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22	Male	No	English	
es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22	Male	No	English	
es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	23	Male	No	English	

Figure 4.22. Filtering data columns to narrow number of rows.

The screenshot shows the CAST Administration interface. On the left, there is a sidebar with navigation links: Approve Items, Email CAST Users, View Algorithm Tests!!!, Demographics, Hindi Administrator, Arabic Item Review, Research Links, and Hindi Instructor Review. The main area displays a table of learner data. The table has columns: Language, Test Date, Registration Code, Test Status, Rated?, Age, Gender, Is language often spoken in your home?, and Is it your language. The table is filtered by the letter 'c' in the Language column, showing three rows of data.

Language	Test Date	Registration Code	Test Status	Rated?	Age	Gender	Is language often spoken in your home?	Is it your language?
ct	2007-11-06 14:46:26.0	BCA_F07	Complete	Rated	21	Male	Yes	Chinese
ct	2007-11-06 14:46:26.0	BCA_F07	Complete	Rated	21	Male	Yes	Chinese
ct	2007-11-06 14:42:05.0	BCA_F07	Complete	Rated	20	Male	Yes	Mandarin

Figure 4.23. Filters column by one letter of the filter criteria.

The screenshot shows the CAST Administration interface with the same sidebar as Figure 4.23. The main area displays a table of learner data. The table is filtered by the value '20' in the Age column, showing five rows of data. The table has columns: Language, Test Date, Registration Code, Test Status, Rated?, Age, Gender, Is language often spoken in your home?, and Is it your language.

Language	Test Date	Registration Code	Test Status	Rated?	Age	Gender	Is language often spoken in your home?	Is it your language?
ct	2007-11-06 14:42:05.0	BCA_F07	Complete	Rated	20	Male	Yes	Mandarin
ct	2007-11-06 14:42:05.0	BCA_F07	Complete	Rated	20	Male	Yes	Mandarin
ct	2007-11-06 14:41:44.0	BCA_F07	Complete	Rated	20	Female	Yes	chinese
ct	2007-11-06 14:41:44.0	BCA_F07	Complete	Rated	20	Female	Yes	chinese

Figure 4.24. Filters column by two letters of filter criteria.

**CAST Administration**

View DA Materials Template Script  
View DA Materials Situation Script  
View CAST Resources  
Development Wiki

**Super User Functions**

Internationalize CAST  
Create New Org Code  
Approve Test Takers  
Add A New Administrator  
Arm/Disarm Questions  
User/Admin Control System  
Review Test Status  
Review Ratings  
Review Reported Concerns/Test Issues  
Approve Administrative User  
Add Language to DA Databases  
Add Language  
Add Context  
Add Function  
Edit Items  
Approve Items  
Email CAST Users  
View Algorithm Test!!!  
Demographics

**Hindi Administrator**

Please review translations and edit regularly and determine when an item is ready to be used in the CAST test.

**CAST Pre-Survey**

☐ Test ID  
☐ Name  
☒ Language  
☒ Test Date  
☒ Registration Code  
☒ Test Status  
☒ Rated  
☒ Age  
☒ Gender

☒ Language often spoken in your home  
☒ Your native language?  
☒ live in area - language was predominant  
☒ language of instruction (primary)  
☒ language of instruction (secondary)  
☒ Taken a "Study Abroad" course  
☒ highest level of education  
☒ Global Ratings

**CAST Post-Survey**

☒ Current TESTING course number  
☒ Years of formal language study  
☒ Residence length  
☒ Understood what I was expected to do  
☒ Instructions for giving responses - clear responses  
☒ Had enough time to prepare my responses  
☒ Had enough time to say what I wanted to say  
☒ Was asked to talk about a variety of topics

☒ Give responses that reflect ability to speak  
☒ Felt comfortable giving my responses  
☒ Enjoyed taking this test  
☒ Items could be improved by  
☒ TESTING program ran smoothly  
☒ Fidelity of the audio  
☒ Volume level of the audio  
☒ Listening to test instructions  
☒ Procedure for recording responses

☒ Recorded voice properly  
☒ The test delivery software could be improved by  
☒ Test Directions  
☒ Timed Pauses  
☒ Nervousness During the Test  
☒ Ideas for Test Improvement  
☒ Any Other Comments  
☒ get Accurate ideas of responses - what extent:

**Learner's Data:**

Language	Test Date	Registration Code	Test Status	Rated?	Age	Gender	Is language often spoken in your home?	Is it your language?
ct					20	female		
ct	2007-11-06 14:41:44.0	BCA_F07	Complete	Rated	20	Female	Yes	chinese
ct	2007-11-06 14:41:44.0	BCA_F07	Complete	Rated	20	Female	Yes	chinese

**Figure 4.25. Filters data columns by multiple filter criteria.**

## **CHAPTER 5**

### **TESTING**

Software testing determines whether a system, as implemented, satisfies its formal stated system requirements. A test case is a set of conditions or variables under which it is determined whether an application or software system is working correctly or not.

Software testing can be stated as the process of validating and verifying to make sure that a software program/application/product meets the requirements that directed its design and development, works as expected and can be implemented with the same characteristics [16].

Unit testing validates the correct behaviour of every minimal executable segment of the application code.

Sub-system/system testing validates the correct interaction between all sub-systems of the system.

Integration testing validates that the system meets stated requirements when individual software modules are combined and tested as a group. Different students were working on different modules of CAST at the same time, so integration testing was carried out to make sure that once all the students deploy their modules to the server, the individual applications still work fine. For this application, the login functionality was caching the user and was not detecting the different types of users logging in. As a result, through testing of this application was carried and some more code was added to admin class to fix the issue. Using Java Random Number Generator functionality, a random number is generated per each user login to avoid user login caching.

The complete set of test descriptions for this project is as follows in Table 5.1.

**Table 5.1. Testing Results**

<b>ID</b>	<b>Test Case</b>	<b>Expected Result</b>	<b>Action</b>	<b>Result</b>
1	Cast Pre-Survey Data Link	Creates Excel sheet for Cast Pre-Survey Data	Click on Cast Pre-Survey Data link	Pass
2	Cast Post-Survey Data Link	Creates Excel sheet for Cast Post- Survey Data	Click on Cast Pre-Survey Data link	Pass
3	Cast Post Pre Survey Data Link	Creates Excel sheet for Cast Post and Pre Survey Data	Click on Cast Post Pre-Survey Data link	Pass
4	Demographics Link	Open up the main screen of the Application	<ol style="list-style-type: none"> <li>1. Go to the Login Screen.</li> <li>2. Sign up as user/administrator</li> <li>3. Click on the Demographics link in the left column.</li> </ol>	Pass
5	Data Grid Column Filter	When we start typing “Female” in Gender column, it gets filtered and shows results for gender “Female”	Type “F” in Gender Column Filter	Pass
6	Position of Pop-up Menu	When user users a right click on the data grid, a quick menu pops up. The quick menu’s location should be limited to the grid context.	Right click on the data grid.	Pass
7	Expected Columns in Excel Sheet	When user clicks on the Post-Pre Survey link, the data from both post and pre survey is correlated and presented in the same workable sheet.	Click on the Post Pre Survey Data link; compare the columns with the original data.	Pass

(table continues)



**Table 5.1. (continued)**

<b>ID</b>	<b>Test Case</b>	<b>Expected Result</b>	<b>Action</b>	<b>Result</b>
8	Make personal data anonymous for researchers other than the administrators	User Information is visible only to the Administrators and for other researchers the personal information of the test-takers should be anonymous.	Sign is a researcher/user. Click on the Demographics link. The name column will read “Anonymous”. Also right click on the Learner’s data grid and select View User Info. All the information will appear as “Anonymous”.	Pass
9.	Linking checkboxes to the research data columns	When the user checks or un-checks a column name, only it’s corresponding data grid column should be shown or hidden respectively.	Uncheck box Name in the categories. The column name will be hidden. Check the box Name in the categories and the column Name will be hidden.	Pass
10.	Verifying Database Results	The title of the data grid columns should match the data in the respective column.	Make sure the data in the data grid is consistent with the data column titles	Pass
11.	Browser Independence -Checkboxes overflowing out of the bounding box	Checkboxes should be bounded	Open the Application with different browser types and the application should appear same in all the browsers.	Pass

## **CHAPTER 6**

### **CONCLUSION AND FUTURE WORK**

#### **6.1 CONCLUSION**

This section lists the evaluation of the application by the LARC department and the project's accomplishments:

- Application is successfully to the main server and is well tested by Mike Pasamonik the Software Engineer at LARC.
- Created an easy-to-use web interface to the learner's database.
- Preserve confidentiality of learner's information by limiting user's access to data. It was achieved by creating different levels of users such as administrators and users.
- The web interface helps researcher create reports in form of workable Excel sheets for the Pre-survey, Post-survey and both together.
- The learner's data should be easy to sort and filter.
- The filter can be applied to many categories at the same time and thus help the researcher create various different patterns of data.
- Limit the data to be displayed on the screen by giving user the option to check and uncheck the search criteria.
- Includes spoken responses from French, German, Japanese, Spanish and other language learners in single web interface.
- The application is searchable by a number of parameters, including school level, class level, proficiency level, and immersion or heritage status
- A user manual is created (Appendix C) to make sure users can utilize this application with less difficulty.

#### **6.2 FUTURE WORK**

As a part of the future work, the following improvements can be made to the application:

1. Implement a method in the application to export the research data after applying different sorting categories.
2. Implement functionality in the application to generate graphs using the results from the existing report results.

## BIBLIOGRAPHY

- [1] Foreign Language Resource Center. About LRCs, 2010. <http://nflrc.msu.edu/>, accessed Sept. 27, 2011.
- [2] Language Acquisition Resource Center. About LARC, 2011. <http://larc.sdsu.edu/>, accessed Sept. 27, 2011.
- [3] The University of Texas at Austin. Spanish Corpus Proficiency Level Training, n.d. <http://www.laits.utexas.edu/spt/intro>, accessed Sept. 27, 2011.
- [4] Language Acquisition Resource Center. Computer Assisted Screening Tool, n.d. <http://cast.sdsu.edu/>, accessed Sept. 27, 2011.
- [5] University of Oregon. About CASLS, 2010. <http://casls.uoregon.edu/pages/about/index.php>, accessed Sept. 27, 2011.
- [6] University of Oregon. Center for Applied Second Language Studies, n.d. <http://casls.uoregon.edu>, accessed Sept. 27, 2011.
- [7] Harvard University. Department of Linguistics Research, 2007. <http://www.fas.harvard.edu/~herpro/research.htm>, accessed Sept. 27, 2011.
- [8] TalkBank.org. TalkBank, n.d. <http://talkbank.org/>, accessed Oct. 6, 2010.
- [9] British Association for Applied Linguistics. BAAL News, 2010. <http://www.baal.org.uk/>, accessed Sept. 27, 2011.
- [10] LSA Ethics. Conference on Language Ethics, 2011. <http://lsaethics.wordpress.com/>, accessed Sept. 27, 2011.
- [11] UCLA Center. Confidentiality and Informed Consent, 2008. [smhp.psych.ucla.edu/pdfdocs/confid/confid.pdf](http://smhp.psych.ucla.edu/pdfdocs/confid/confid.pdf), accessed Sept. 27, 2011
- [12] Oracle. Java Servlet Technology, n.d. <http://www.oracle.com/technetwork/java/javaee/servlet/index.html>, accessed Sept. 27, 2011.
- [13] jQuery. jQuery Project, 2010. <http://jQuery.com/>, accessed Oct. 5, 2010.
- [14] w3schools.com. JavaScript Tutorial, n.d. <http://www.w3schools.com/js/default.asp>, accessed Sept. 27, 2011.
- [15] JExcelApi. Java Excel API, n.d. <http://JExcelapi.sourceforge.net/>, accessed Sept. 27, 2011.
- [16] Wikipedia, Software Testing, n.d. [http://en.wikipedia.org/wiki/Software\\_testing/](http://en.wikipedia.org/wiki/Software_testing/), accessed Sept. 27, 2011.

**APPENDIX A**

**DIAGNOSTIC ASSESSMENT USING CAST**

**CONSENT FORM**

## Diagnostic Assessment using the Computer Assisted Screening Tool (CAST)

### Consent Form

We are asking you to participate in a study being conducted by Dr. Mary Ann Lyman-Hager of the Language Acquisition Research Center (LARC) at San Diego State University. The Institutional Review Board at SDSU has approved this study. The purpose of this study is to obtain information about what foreign language learners' speech sounds like at various levels of language proficiency. Another purpose of our research is to develop a framework to give language learners feedback on their language performance. Anyone 18 years and older may participate. Your participation, including the pre-test questionnaire and the on-line speaking test, is very important to us and should last less than one hour. Depending on where you participate, you may or may not receive financial compensation. If you are participating at an institution where you do receive an incentive, you may receive up to [ ].

In a pre-test questionnaire, we will ask information about your language background and your experiences using the language. In this questionnaire there are no right and wrong answers. You will next be given an on-line speaking test and will record your responses to specific questions. Questions will consist of two parts. The first part, describing a situation, is both written and recorded. The second part is an audio prompt. Your recorded answers will be sent to an evaluator for analysis. You may receive some basic feedback on your language skills.

Data from the speaking test will also be valuable to linguistic researchers who examine speech patterns or structures used by language learners at different levels. Your data may be used in linguistic research following this study. Your identity will be coded so that the researchers will be unable to associate your voice with your identity.

There will be several pages of instruction to the oral test. You may complete the oral test at your leisure, taking all the time you need and responding as completely as you can. Please answer each question verbally by speaking clearly into the microphone. Your oral response will be sent to the server.


All responses to these questions are **TOTALLY CONFIDENTIAL**.

In summary, the study involves completing a number of questions that ask you about your language experiences; the answers you provide will be kept confidential. You are free to withdraw your consent at any time. The speaking data you provide will be evaluated. These data may be used for future research on language learning, but your identity will be hidden from researchers. Your completion of the questionnaire survey and the oral test are completely voluntary.

We encourage you to contact us at 619-594-5445 if you feel you would like to discuss the survey, your score or the speaking test. If you have concerns about your treatment as a subject, please feel free to contact the SDSU Institutional Review board at 619-594-6622. Thank you for your participation.

\_\_\_\_\_ I consent

\_\_\_\_\_ I do not consent

 <b>SAN DIEGO STATE UNIVERSITY</b>	<b>Institutional Review Board</b> Approval Expires: 8/22/2012 Study Number: 3555
--	--

**APPENDIX B**

**DIAGNOSTIC ASSESSMENT (EXTENSION OF  
CAST: RENEWAL APPROVAL FOR  
PROTOCOL #02-09-278)**

Graduate and Research Affairs  
Division of Research Affairs  
San Diego State University  
5500 Campanile Drive  
San Diego, CA 92182-8220  
Phone: 619-594-5938  
Fax: 619-594-4109

**NEXT REPORT OF PROGRESS DUE 7/22/2012**

See below for further information on Protocol Maintenance submissions and  
NEW IRB REQUIREMENTS!

Faculty Researcher: Mary Ann Lyman-Hager, Ph.D.  
Department: European Studies  
Title: Diagnostic Assessment (Extension of CAST: renewal approval for Protocol #02-08-278)  
vIRB Number: 3333  
Grants/Contracts: 52543, a,b,c (federal, other)  
Regulatory Determination  
Risk Level: Minimal  
Continuation approved per 45 CFR 46.110, Categories 6 & 7

The SDSU Institutional Review Board approved the project referenced for continuation on August 22, 2011 in accordance with SDSU's Assurance and federal requirements pertaining to human subjects protections within the Code of Federal Regulations (45 CFR 46; 21 CFR 30). This approval only applies to the analysis of data collected prior to 8/11/2011 and the recruitment of human subjects and data collection after 8/22/2011 and analysis of that data based on procedures described in your protocol and the information you provided within your report of progress. Approval carries with it the understanding that you will contact the IRB to obtain authorization to implement any proposed changes to the protocols, to document a change in your affiliation with SDSU (student, faculty or staff), and/or to report the completion of data analysis. Project approval expires August 22, 2012.

**Please note your expiration date.** To request continued recruitment, data collection and/or data analyses, a Report of Progress must be submitted prior to the expiration date of your study. A lapse in approval requires that all research with human subjects be suspended until approval is obtained and may result in a temporary hold on funds, if your study is funded. The investigator will be out of compliance with federal regulation and university policy if human subjects continue to be involved in this project without a valid IRB approval.

As your study requires the use of consent document(s), please note that the re-stamped form(s) has been uploaded to your protocol file within the VIBR system, within the "Supporting Documents" section. This document bears the IRB's stamp of approval. Print a copy of this stamped form to use when documenting informed consent from research participants. *Changes*

may not be made to the consent documents without prior review and approval of the IRB. You are required to keep signed copies of the consent documents for three years after your project has been completed or terminated.

■ CastConsentFINAL, stamped to expire 8-22-2012.pdf

a) For studies requiring consent translation: The SDSU Institutional Review Board (IRB) does not verify the accuracy of the translated document. IRB approval of this document for use in subject recruitment is based on your assurance that the translated document reflects the content of the IRB approved English version of the document.

b) If recruitment will take place through an outside agency or organization, confirm with that institution that you have permission to conduct the study prior to initiation of any study activities.

c) Approval is contingent upon the completion of the SDSU human subjects tutorial (found at: <http://www-rohan.sdsu.edu/~graj/login.php>) by all members of the research team. This certification must be renewed every 2 years.

d) If any changes to your study are planned, you must submit a modification request and receive IRB approval prior to the implementation of study changes. To submit a modification request, access the protocol via the WebPortal, on the protocol Main Page, you will need to click on "Modifications" under Protocol Maintenance and enter a report. Once you have filled in your responses on the report form, click "submit".

- To document your modification in detail, access your currently approved protocol in the "Full Document Viewer." Copy and paste the document into Word and use "track changes" to document revisions to your protocol. Save the file (Name\_Modification\_Date) and upload it to your protocol file. When approved by the IRB, this document will be the current version of your approved protocol.
- If a change to the approved consent form(s) or other uploaded document(s) is being requested, changes must be documented using the "track changes" feature in Word. Upload the revised form to your IRB protocol file. This form will be reviewed by the IRB. If you do not have a copy of your approved consent form in a Word format, request a copy from the IRB office.

e) The SDSU IRB requires investigators to report any problems that arise during the course of an IRB approved research study. *Serious adverse events or unanticipated problems that are life-threatening or have resulted in serious injury or death must be reported to the IRB immediately whenever possible or within at least 48 hours from the onset of the incident. All other problems must be reported to the SDSU*





Graduate and Research Affairs  
Division of Research Affairs  
San Diego State University  
5500 Campanile Drive  
San Diego, CA 92182-8220  
Phone: 619-594-3938  
Fax: 619-594-4109

IRB within 3 days. To complete and submit an adverse event report, go to the Protocol Main Menu, click on "Adverse Events" under "Protocol Maintenance" and follow the instructions. For more information and consultation, contact the IRB office directly via Email at: [IRB@mail.sdsu.edu](mailto:IRB@mail.sdsu.edu) or telephone: 619-594-6622, Monday through Friday from 8:00AM to 4:00PM.

f) To submit a request to extend IRB approval, log in to your WebPortal account and access the protocol. On the protocol Main Page, click on "Progress Reports" under Protocol Maintenance and enter a report. Once you have filled in your responses on the report form, click "submit". You should receive an automated email verifying IRB receipt of your Report of Progress.

**NEW REQUIREMENT!** Within the description box of the Report of Progress form, indicate which, if any, consent form(s) you are requesting to renew. Refer to the Consent Form Development section of the protocol and provide the IRB with the specific file names and date(s) of upload of the consent document(s) you are requesting to renew.

For questions related to this correspondence, please contact the IRB office ((619) 594-6622 or e-mail [irb@mail.sdsu.edu](mailto:irb@mail.sdsu.edu)). To access IRB review application materials, SDSU's Assurance, the 45 CFR 46, the Belmont Report, and/or any other relevant policies and guidelines related to the involvement of human subjects in research, please visit the IRB web site at <http://gra.sdsu.edu/research.php>.

**Graduate Students:** This notification may be used as documentation to register in Thesis 799A. Attach a hard copy of this notice to your Appointment of Thesis/Project Committee form prior to submitting the completed form to Graduate and Research Affairs - Student Services Division.

Sincerely,

Jeanne Nichols  
Chair, Institutional Review Board

Amy McDaniel  
Regulatory Compliance Analyst

Brianne Larsen-Mongson  
Regulatory Compliance Analyst

Choya Washington  
Regulatory Compliance Analyst

**APPENDIX C**  
**USER MANUAL**

## User Manual

### Index

1. GENERAL INFORMATION .....	42
1.1 Application Overview .....	42
1.2 Point of Contact .....	42
2. GETTING STARTED .....	43
2.1 User Login Process .....	43
2.2 Download Excel Sheet Reports .....	43
2.3 Filter Learner's Data .....	45
2.4 Show or Hide Learner's Data .....	47
2.5 Using Context Menu .....	47

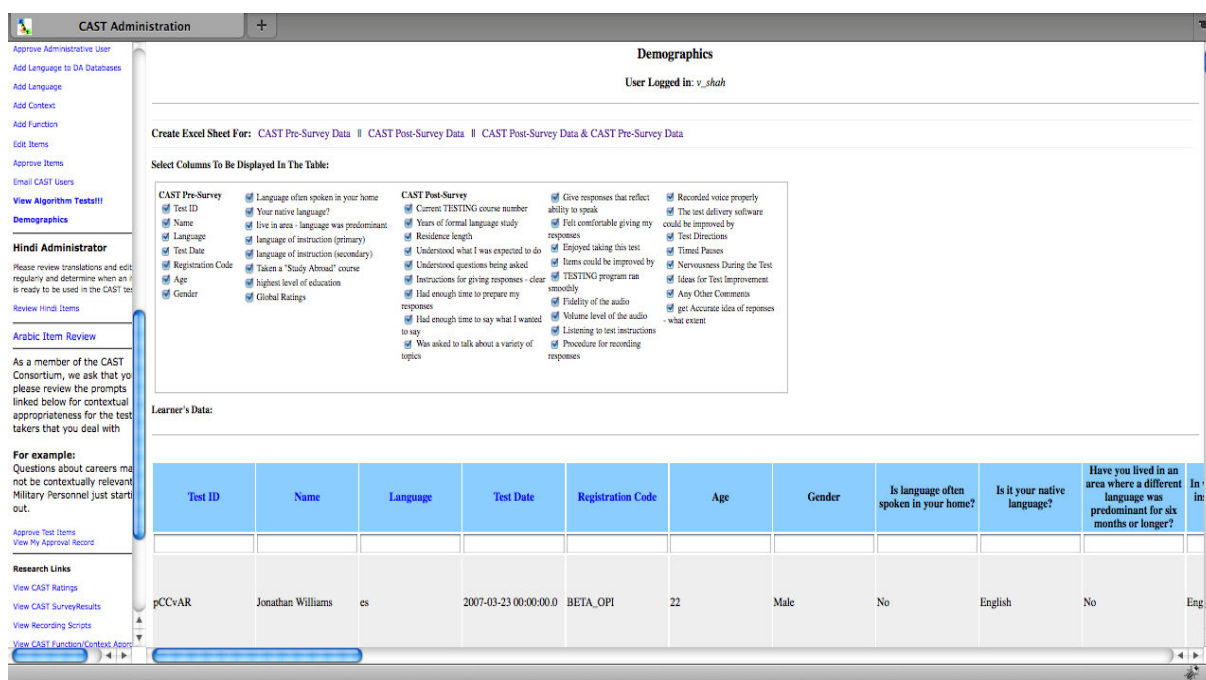
## 1 General Information

This section describes the general information about the application. It also describes the application overview and the point of contact in case of any questions about the application.

### 1.1 Application Overview

The Application has the following features:

- **Create Excel Sheet** – The “Create Excel Sheet For” section on the screen as shown in the Figure 1 is used to generate MS Excel reports for Pre-survey, Post-survey and both pre and post survey data.
- **Learner’s Data Table** – The “Learner’s Data” section on the screen as shown in the Figure 1 is used to display all the learner data together in a table format.
- **Column Filter** – The text box above each column in the Learners’ data table as shown in the Figure 1 is used as a search box to filter through the column and narrow down the number of rows to match the filter categories entered in this search box.



**Figure 1: Main Application Screen**

### 1.2 Point of Contact

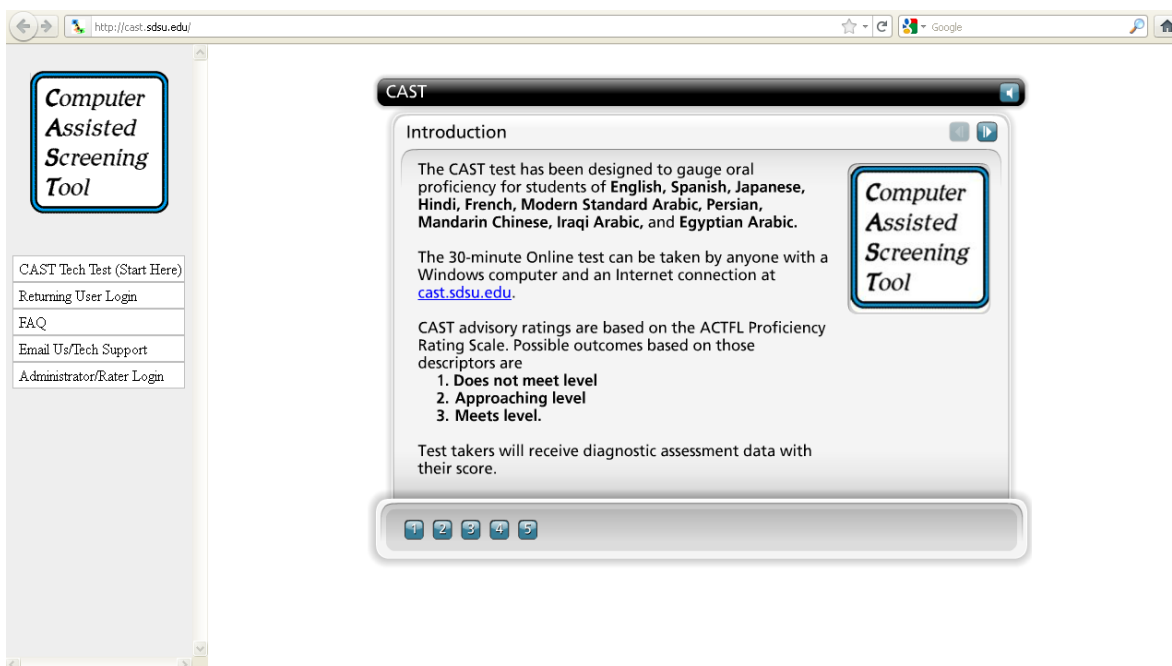
This application is created for the linguistic research work that is carried out at the Language Acquisition Resource Center. In case of any questions, contact LARC at San Diego State University.

## 2 Getting Started

This section describes the guidelines to use this application by listing the steps to use each feature of the application.

### 2.1 User Login Process

This section will guide you through the User login process to access the demographics application. First, go to the page <http://cast.sdsu.edu>. Click on the Administrator/Rater login as shown in the Figure 2.

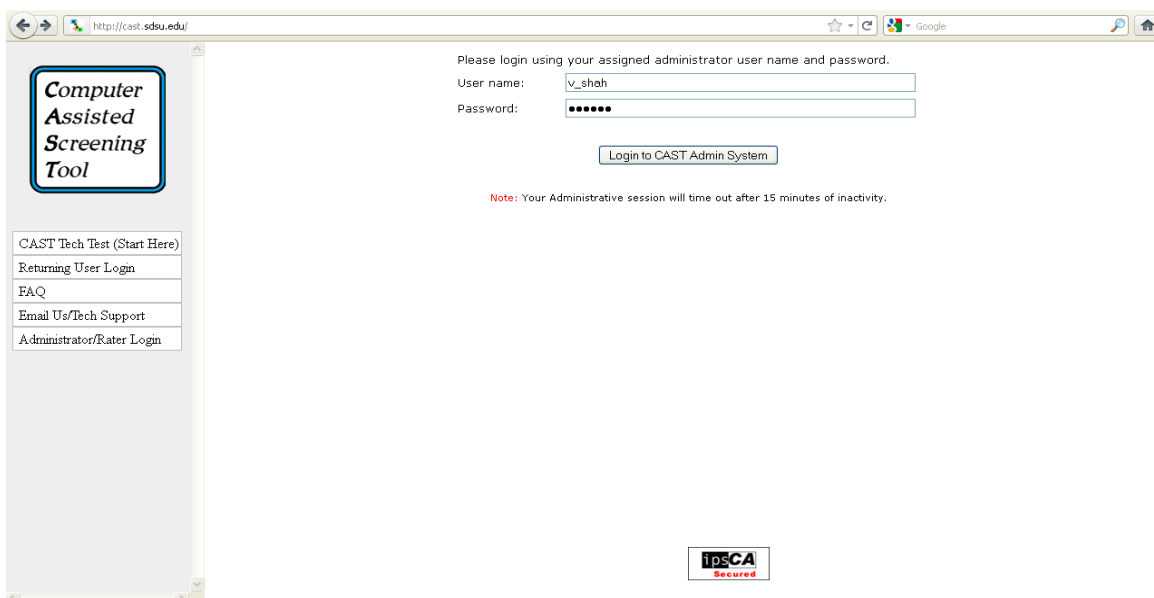


**Figure 2: CAST Screen**

Enter User ID and password as shown in the Figure 3. Contact the administrator if you do not have required login credentials. Scroll down and click on the “Demographics – Survey results” link to go to the main application.

### 2.2 Generate Excel Reports

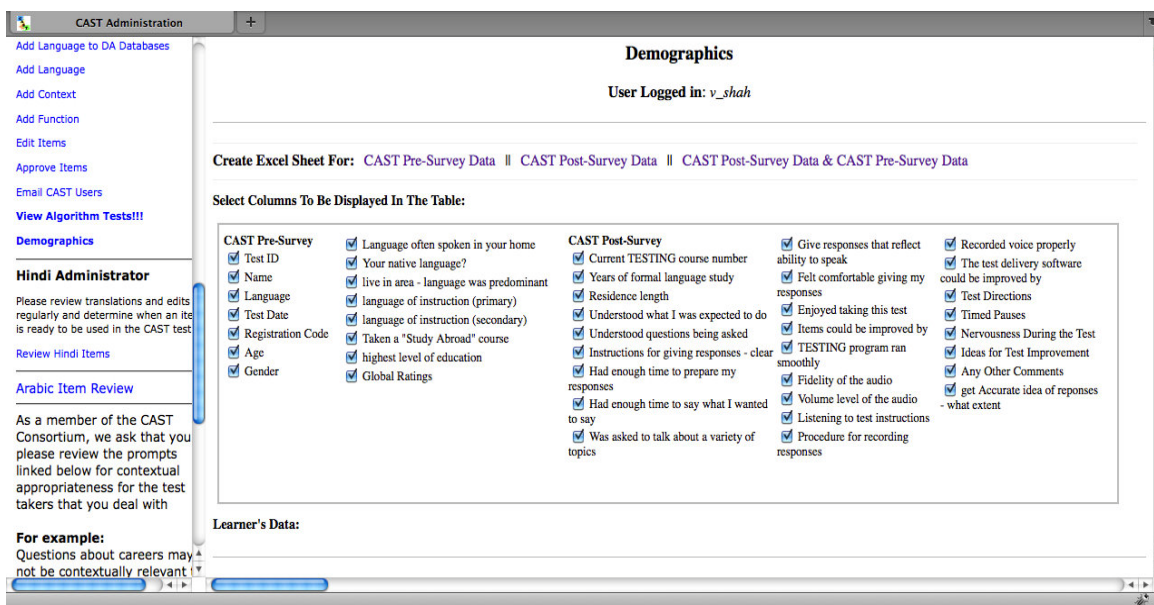
The section “Generate Excel Report” highlighted in red in the figure below allows you to create Pre-Survey, Post-Survey and Pre-Survey/Post-Survey data reports, which enables you to work with the survey data in Excel format. These reports also include the Global Ratings of the language learner, which enables you to relate their survey results with their performance on the test.



**Figure 3: Login Screen**

## 2.2.1 Create Pre-Survey Report

Click on the link “Pre-Survey Report” as shown in Figure 4. It will open up an Open/Save Dialog box. Click on “*open*” to open up the Excel report or click on *save* to save it at a selected location. The Excel report will be generated in the format as shown in Figure 4.



**Figure 4: Generate Excel Sheet Report**

## 2.2.2 Create Post-Survey Report

Click on the link “Post-Survey Report” as shown in Figure 4. It will open up an Open/Save Dialog box. Click on “*open*” to open up the Excel report or click on save to save it at a selected location. This link will create an excel sheet report similar to the one shown in the Figure 5.

Age	Gender	Is language often spoken	Is it your native language	Have you lived in an area	In what language was it	Have you taken a Student	What is the highest level of education	Global Ratings
22	Male	No	English	No	English	English No	High School	full
22	Male	No	English	No	English	English No	High School	full
22	Male	No	English	No	English	English No	High School	full
23	Male	No	English	Yes	English	English No	High School	full
24	Female	No	English	No	English	English Yes	Associate Degree	full
24	Male	No	English	No	English	English No	Associate Degree	full
21	Male	No	English	Yes	English	English No	High School	full
21	Male	No	English	Yes	English	English No	Associate Degree	full
21	Male	No	English	Yes	English	English No	High School	full
23	Female	No	English	Yes	english	English No	High School	unrate
25	Female	No	English	Yes	English	English No	Associate Degree	approach
23	Male	No	English	Yes	English	English No	High School	full
18	Female	No	English	No	English	English No	High School	fails
23	Male	No	English	Yes	English	English Yes	High School	full
21	Female	No	English	No	English	English Yes	High School	null
21	Male	No	English	Yes	English	English No	High School	null
19	Male	No	English	No	English	English No	High School	null
23	Male	No	English	Yes	English	English Yes	Associate Degree	null
19	Female	No	English	No	english	English No	High School	approach
21	Male	No	English	Yes	English	English No	High School	full
20	Female	No	English	No	English	English No	High School	full
24	Male	No	English	Yes	English	English Yes	Associate Degree	full
23	Female	No	English	Yes	English	English No	High School	approach
23	Female	No	English	Yes	English	English No	High School	approach
19	Female	Yes	English	No	English	English No	High School	fails
19	Female	No	English	No	English	English No	High School	fails
19	Female	No	English	No	English	English No	High School	fails
19	Female	No	English	No	English	English No	High School	fails
22	Male	No	English	Yes	English	English No	High School	approach
22	Male	No	English	Yes	English	English No	High School	approach
22	Male	No	English	Yes	English	English No	High School	approach

Figure 5: Sample Excel Report

## 2.2.3 Create Pre/Post Survey Report

Click on the link “Pre/Post Survey Report” as shown in Figure 4. It will open up an Open/Save Dialog box. Click on “*open*” to open up the Excel report or click on save to save it at a selected location. This link will create an Excel sheet report similar to the one shown in the Figure 5.

## 2.3 Show/Hide Column's of Learner's data

The checkboxes under the “Select columns to be displayed in the table” section are mapped to the columns in the learner's data table. If you uncheck a checkbox, the column mapped to it in the table will also be hidden. If you want to column again – uncheck the checkbox mapped to it. For example, as shown in the Figure 6, the Test ID box is selected, so

CAST Administration

Demographics

User Logged in: v\_ah

Create Excel Sheet For: CAST Pre-Survey Data | CAST Post-Survey Data | CAST Post-Survey Data & CAST Pre-Survey Data

Select Columns To Be Displayed In The Table:

**CAST Pre-Survey**

- ☒ Test ID
- ☒ Name
- ☒ Language
- ☒ Test Date
- ☒ Registration Code
- ☒ Age
- ☒ Gender
- ☒ Language often spoken in your home?
- ☒ Your native language?
- ☒ Live in area - language was predominant
- ☒ Language of instruction (primary)
- ☒ Language of instruction (secondary)
- ☒ Taken a "Study Abroad" course
- ☒ Highest level of education
- ☒ Global Ratings

**CAST Post-Survey**

- ☒ Current TESTING course number
- ☒ Years of formal language study
- ☒ Residence length
- ☒ Understood what I was expected to do
- ☒ Understood questions being asked
- ☒ Instructions for giving responses - clear
- ☒ Had enough time to prepare my responses
- ☒ Had enough time to say what I wanted to say
- ☒ Was asked to talk about a variety of topics
- ☒ Give responses that reflect ability to speak
- ☒ Felt comfortable giving my responses
- ☒ Enjoyed taking this test
- ☒ Items could be improved by
- ☒ TESTING program ran smoothly
- ☒ Fidelity of the audio
- ☒ Volume level of the audio
- ☒ Listening to test instructions
- ☒ Procedure for recording responses
- ☒ Recorded voice properly
- ☒ The test delivery software could be improved by
- ☒ Test Directions
- ☒ Timed Pauses
- ☒ Nervousness During the Test
- ☒ Ideas for Test Improvement
- ☒ Any Other Comments
- ☒ get Accurate idea of responses - what extent

**Learner's Data:**

Test ID	Name	Language	Test Date	Registration Code	Age	Gender	Is language often spoken in your home?	Is it your native language?	Have you lived in an area where a different language was predominant for six months or longer?	In
	Jonathan Williams	es	2007-03-23 00:00:00.0	BETA_OPI	22	Male	No	English	No	Eng

**Figure 6: Check box to hide the appropriate column – Test ID column is visible**

the Test ID column is visible. If the Test ID box is unchecked, the Test ID column in the table will be hidden as shown in the Figure 7.

CAST Administration

Select Columns To Be Displayed In The Table:

**CAST Pre-Survey**

- ☐ Test ID
- ☒ Name
- ☒ Language
- ☒ Test Date
- ☒ Registration Code
- ☒ Age
- ☒ Gender
- ☒ Language often spoken in your home?
- ☒ Your native language?
- ☒ Live in area - language was predominant
- ☒ Language of instruction (primary)
- ☒ Language of instruction (secondary)
- ☒ Taken a "Study Abroad" course
- ☒ Highest level of education
- ☒ Global Ratings

**CAST Post-Survey**

- ☒ Current TESTING course number
- ☒ Years of formal language study
- ☒ Residence length
- ☒ Understood what I was expected to do
- ☒ Understood questions being asked
- ☒ Instructions for giving responses - clear
- ☒ Had enough time to prepare my responses
- ☒ Had enough time to say what I wanted to say
- ☒ Was asked to talk about a variety of topics
- ☒ Give responses that reflect ability to speak
- ☒ Felt comfortable giving my responses
- ☒ Enjoyed taking this test
- ☒ Items could be improved by
- ☒ TESTING program ran smoothly
- ☒ Fidelity of the audio
- ☒ Volume level of the audio
- ☒ Listening to test instructions
- ☒ Procedure for recording responses
- ☒ Recorded voice properly
- ☒ The test delivery software could be improved by
- ☒ Test Directions
- ☒ Timed Pauses
- ☒ Nervousness During the Test
- ☒ Ideas for Test Improvement
- ☒ Any Other Comments
- ☒ get Accurate idea of responses - what extent

**Learner's Data:**

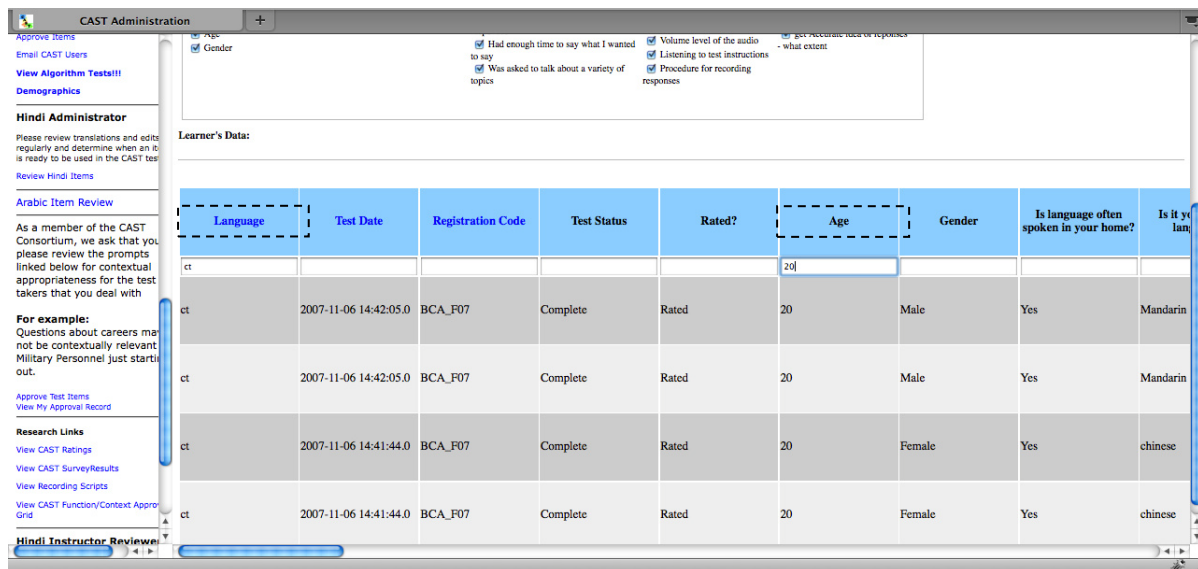
Name	Language	Test Date	Registration Code	Age	Gender	Is language often spoken in your home?	Is i
Jonathan Williams	es	2007-03-23 00:00:00.0	BETA_OPI	22	Male	No	Eng

**Figure 7: Test ID Column is hidden**



## 2.4 Filter Learner's Data Columns

The textbox above each column as shown with a dotted line in Figure 8 acts as search box. As the user starts typing in the box, the column is filtered accordingly. The user can filter multiple columns at the same time. For example, when the user enters two in the Age column and c in language column filter the rows in the column “Age” are narrowed to the ones with value 2.



**Figure 8: Column Filter functionality**

## 2.5 Using Popup Menu

When the user hits the “right click” on the Learner’s data table a context menu pops up as shown in the Figure 9. This context menu has three options as shown in the figure below:

- View User Info – When the user clicks on View User Info, he can see the user’s contact information.
- Sort Users By – User can use this option to implement the functionality to sort the data by Test ID, Name, or Language.
- View User’s Exam – User can use this option to view the audio recordings of the language learners, global ratings, and rater comments.

CAST Administration

User/Admin Control System

Review Test Status

Review Ratings

Review Reported Concerns/Test Issues

Approve Administrative User

Add Language to DA Databases

Add Language

Add Context

Add Function

Edit Items

Approve Items

Email CAST Users

View Algorithm Tests!!!

Demographics

**Spanish Administrator**

Please review translations and edit items regularly and determine when an item is ready to be used in the CAST test.

Review Spanish Items

Arabic Item Review

As a member of the CAST Consortium, we ask that you please review the prompts linked below for contextual appropriateness for the test.

Listening to test instructions

Was asked to talk about a variety of topics

Procedure for recording responses

**Learner's Data:**

Test ID	Name	Language	Test Date	Registration Code	Test Status	Rated?	
pCCvAR	Anonymous	es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22
pCCvAR	Anonymous		2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	22
CjgDNn	Anonymous	es	2007-03-23 00:00:00.0	BETA_OPI	Complete	Rated	23

**QUICK MENU**

- View User Info
- Sort Users By
- View User's Exam

**Figure 9: Quick Menu**